Universidade de Aveiro
2011
Departamento de Química

João Eduardo A. Rodrigues

Estudos Metabonómicos sobre a Qualidade e Estabilidade da Cerveja

Metabonomic Studies to Assess Beer Quality and Stability

Universidade de Aveiro Departamento de Química
2011

João Eduardo A. Rodrigues

Estudos Metabonómicos sobre a Qualidade e Estabilidade da Cerveja

Metabonomic Studies to Assess Beer Quality and Stability

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Doutorado em Química, realizada sob a orientação científica da Doutora Ana Maria Gil, Professora Associada com Agregação do CICECO, Departamento de Química da Universidade de Aveiro e do Doutor António de Sousa Barros, Investigador Auxiliar do QOPNA, Departamento de Química da Universidade de Aveiro.

Apoio financeiro da FCT através da bolsa FCT SFRH/BD/31056/2006.

o júri

presidente

Professor Doutor José Manuel Lopes da Silva Moreira
Professor Catedrático, Departamento de Ciências Sociais, Políticas e do Território, Universidade do Aveiro, Portugal

Doutora Ana Maria Pissarra Coelho Gil
Professora associada com agregação, CICECO-Departamento de Química, Universidade do Aveiro, Portugal

Doutor António de Sousa Barros
Investigador auxiliar, QOPNA-Departamento de Química, Universidade de Aveiro, Portugal

Professor Doutor Douglas Rutledge
Full professor of Analytical Chemistry and Director of Laboratory of Analytical Chemistry at Ingénierie Analytique pour la Qualité des Aliments (IAQA), Laboratoire de Chimie Analytique, AgroParisTech, France

Doutor António César Ferreira
Investigador da Escola Superior de Biotecnologia da Universidade Católica Portuguesa, Porto, Portugal, e Full professor of Analytical Chemistry at Institute for Wine Biotechnology and Department of Viticulture and Oenology, Stellenbosh University, South Africa

Mestre Tiago Brandão
Manager de Inovação e Desenvolvimento, Direcção de Marketing, Unicer, Bebidas, SA, Portugal

Acknoledgments

I would like to express my sincere gratitude to my supervisors Dr. Ana M. Gil and Dr. António Barros for giving me the opportunity to work with them in the field of beer metabonomics. I would like to thank them for all the support and guidance during the time of my PhD, allowing me to grow as a researcher.

I specially acknowledge Tiago Brandão, Beatriz Carvalho and Cristina Gonçalves from UNICER, Bebidas de Portugal, for the brewing scientific support and disponibility shown. My acknowledgments also go to Dr. Manfred Spraul, Dr. Ebehhard Humpfer and Dr. Hartmutt Schaffer, from Bruker BioSpin, also for all the technical support and expertise shared in the field of NMR spectroscopy.

I would also like to thank Dr. António C. Ferreira, for its help and useful discussions about the aging studies; Dr. Valdemar Esteves and Guillaume Erny for their help in the quantification studies; and Dr. Eurico Cabrita for his contribution in quantitative NMR studies.

This work would not be performed without the scientific help and friendship always demonstrated for the metabonomics group of University of Aveiro. To I. Duarte, B. Goodfellow, J. Carrola, S. Diaz, G. Graça, I. Lamego, J.Pinto and C. Rocha, as well as to J. Marques and J. Simões, my special thanks. Furthermore, for all the people of the Chemistry Department my gratitude for the companionship.

I wish to acknowledge the Portuguese Foundation for Science and Technology (FCT) for their financial support through the grant SFRH/BD/31056/2006. I also thank the CICECO-Chemistry Department of the University of Aveiro for providing the physical conditions for this work.

Finally, I deeply thank my mother, father and sister for their love and continuous encouragement. Also, I express my sincere thanks to all my family and friends for being always present, their support and friendship help me overcome this journey.

Palavra chave Cerveja, controlo de qualidade, envelhecimento, metabonómica em alimentos, espectroscopia de ressonância magnética nuclear (RMN), cromatografia em fase gasosa-espectrometria de massa (GC-MS), análise multivariada (MVA), fusão de dados, quantificação

Abstract

O trabalho apresentado nesta tese pretendeu aplicar a metodologia usualmente designada por metabonómica, baseada na utilização de métodos de análise multivariada (MVA) para extrair a informação relevante de conjuntos de dados, obtidos por técnicas analíticas, no estudo detalhado de um determinado tipo de cerveja e seu controlo de qualidade. No Capítulo 1, é apresentada uma descrição detalhada sobre cerveja: seu processo de fabrico, principais características e composição típica, estabilidade da cerveja e principais técnicas utilizadas para o estudo da composição da cerveja e do seu controlo de qualidade. Os fundamentos teóricos das técnicas analíticas utilizadas no decorrer da tese, nomeadamente espectroscopia de ressonância magnética nuclear (RMN), cromatografia em fase gasosa-espectrometria de massa (GC-MS) e espectroscopia de infravermelho-médio (MIR), conjuntamente com uma descrição da metodologia de metabonómica estão descritos no Capítulo 2.

No Capítulo 3, está descrita a aplicação de RMN de alta resolução para caracterizar detalhadamente a composição química de um tipo de cerveja. O espectro de RMN de ^{1}H obtido pela análise directa de cerveja apresenta uma elevada complexidade, confirmando a capacidade da espectroscopia de RMN para detecção de uma grande variedade de famílias de compostos, numa só corrida. Experiências de RMN bidimensionais foram realizadas para uma identificação detalhada espectral, tendo-se conseguido identificar cerca de 40 compostos, incluindo álcoois, aminoácidos, ácidos orgânicos, nucleósidos e açúcares. Numa segunda parte do Capítulo 3, a variabilidade composicional da cerveja foi estudada. Para tal, a metodologia de metabonómica foi aplicada aos dados de RMN de ^{1}H (RMN/MVA) para avaliar a variabilidade existente entre cerveja da mesma marca, produzida no mesmo país mas diferindo nas instalações e datas de produção. Diferenças entre instalações e/ou datas de produção foram identificadas, revelando variações composicionais relacionadas com alguns passos do processo, tais como brassagem, fermentação e maturação.

O Capítulo 4 descreve a quantificação de ácidos orgânicos em cerveja por métodos baseados na espectroscopia de RMN. Diferentes métodos quantitativos foram desenvolvidos e comparados, nomeadamente por integração directa dos sinais de RMN (*vs.* referência interna ou *vs.* uma referência externa electrónica, método ERETIC), e por métodos estatísticos quantitativos (usando a regressão parcial em mínimos quadrados (PLS)). Os modelos de PLS foram construídos usando como métodos quantitativos de referência a electroforese capilar com detecção directa e indirecta e ensaios enzimáticos. A comparação dos resultados obtidos pelos modelos de PLS1-RMN que demonstraram melhor poder de previsão com os resultados obtidos pelos métodos de integração directa foi realizada, estando estes últimos em concordância com os obtidos pelos modelos de PLS1, apesar de alguma sobreestimativa ter sido detectada na quantificação dos ácidos málico e pirúvico, assim como uma aparente subestimativa para o ácido cítrico.

Finalmente, no Capitulo 5, desenvolveram-se estudos

metabonómicos para uma melhor compreensão do processo de envelhecimento forçado da cerveja. O processo de envelhecimento forçado (durante 18 dias, a 45 ºC) de um lote de cervejas foi acompanhado por i) RMN, ii) GC-MS e iii) MIR. Aplicando os métodos de MVA a cada um dos conjuntos de dados, observou-se uma clara separação entre dias de envelhecimento para os dados de RMN e de GC-MS, com a identificação de compostos significativamente relacionados com o processo de envelhecimento, nomeadamente 5-hidroximetilfurfural (5-HMF), ácidos orgânicos, ácido γ-amino butírico (GABA), prolina, e a razão de dextrinas lineares/ramificadas (por RMN) e 5-HMF, furfural, succinato de dietilo e fenilacetaldeído (marcadores do envelhecimento da cerveja) e, pela primeira vez, 2,3-dihidro-3,5-dihidroxi-6-metil-4(H)-piran-4-ona (DDMP) e maltoxazina (por GC-MS). Para os dados de MIR, nenhuma tendência de envelhecimento foi identificada nos moldes experimentais investigados. A correlação entre os dados de RMN e GC-MS foi também realizada através dos métodos estatísticos de análise do produto externo (OPA) e por correlação espectral interdomínios (*statistical heterospectroscopy*-SHY), permitindo a identificação de um maior número de compostos (11 compostos, 5 dos quais ainda não estão atribuídos) relacionados directamente com o envelhecimento, originando uma descrição mais completa das variações composicionais que ocorrem durante o processo de envelhecimento. A correlação entre dados sensoriais e os dados analíticos obtidos por RMN e GC-MS também foi avaliada, usando para tal, modelos de regressão de PLS tendo os dados sensoriais como referência, Os resultados obtidos demonstram uma relação entre os dados analíticos e os sensoriais, especialmente para a região aromática do espectro de RMN e para os dados obtidos por GC-MS ($r > 0.89$). Contudo, o poder de previsão de todos os modelos de PLS1 construídos foi relativamente baixo, reflectindo o reduzido número de amostras/provadores utilizados neste estudo.

Keywords Beer, quality control, aging, metabonomics in foods, nuclear magnetic resonance (NMR) spectrosocpy, gas chromatography-mass spectrometry (GC-MS), multivariate analysis (MVA), data fusion, quantification

Abstract

The work reported in this thesis aimed at applying the methodology known as metabonomics to the detailed study of a particular type of beer and its quality control, with basis on the use of multivariate analysis (MVA) to extract meaningful information from given analytical data sets. In Chapter 1, a detailed description of beer is given considering the brewing process, main characteristics and typical composition of beer, beer stability and the commonly used analytical techniques for beer analysis. The fundamentals of the analytical methods employed here, namely nuclear magnetic resonance (NMR) spectroscopy, gas-chromatography-mass spectrometry (GC-MS) and mid-infrared (MIR) spectroscopy, together with the description of the metabonomics methodology are described shortly in Chapter 2.

In Chapter 3, the application of high resolution NMR to characterize the chemical composition of a lager beer is described. The ^{1}H NMR spectrum obtained by direct analysis of beer show a high degree of complexity, confirming the great potential of NMR spectroscopy for the detection of a wide variety of families of compounds, in a single run. Spectral assignment was carried out by 2D NMR, resulting in the identification of about 40 compounds, including alcohols, amino acids, organic acids, nucleosides and sugars. In a second part of Chapter 3, the compositional variability of beer was assessed. For that purpose, metabonomics was applied to ^{1}H NMR data (NMR/MVA) to evaluate beer variability between beers from the same brand (lager), produced nationally but differing in brewing site and date of production. Differences between brewing sites and/or dates were observed, reflecting compositional differences related to particular processing steps, including mashing, fermentation and maturation.

Chapter 4 describes the quantification of organic acids in beer by NMR, using different quantitative methods: direct integration of NMR signals (*vs.* internal reference or *vs.* an external electronic reference, ERETIC method) and by quantitative statistical methods (using the partial least squares (PLS) regression) were developed and compared. PLS1 regression models were built using different quantitative methods as reference: capillary electrophoresis with direct and indirect detection and enzymatic essays. It was found that NMR integration results generally agree with those obtained by the best performance PLS models, although some overestimation for malic and pyruvic acids and an apparent underestimation for citric acid were observed.

Finally, Chapter 5 describes metabonomic studies performed to better understand the forced aging (18 days, at 45 ºC) beer process. The aging process of lager beer was followed by i) NMR, ii) GC-MS, and iii) MIR spectroscopy. MVA methods of each analytical data set revealed clear separation between different aging days for both NMR and GC-MS data, enabling the identification of compounds closely related with the aging process: 5-hydroxymethylfurfural (5-HMF), organic acids, γ-amino butyric acid (GABA), proline and the ratio linear/branched dextrins (NMR domain) and 5-HMF, furfural, diethyl succinate and phenylacetaldehyde (known aging markers) and, for the first time, 2,3-dihydro-3,5-dihydroxy-6-methyl-4(H)-pyran-4-one

(DDMP) and maltoxazine (by GC-MS domain). For MIR/MVA, no aging trend could be measured, the results reflecting the need of further experimental optimizations. Data correlation between NMR and GC-MS data was performed by outer product analysis (OPA) and statistical heterospectroscopy (SHY) methodologies, enabling the identification of further compounds (11 compounds, 5 of each are still unassigned) highly related with the aging process. Data correlation between sensory characteristics and NMR and GC-MS was also assessed through PLS1 regression models using the sensory response as reference. The results obtained showed good relationships between analytical data response and sensory response, particularly for the aromatic region of the NMR spectra and for GC-MS data ($r > 0.89$). However, the prediction power of all built PLS1 regression models was relatively low, possibly reflecting the low number of samples/tasters employed, an aspect to improve in future studies.

Publications of some of the work reported here:

Papers

Erny, G.L.; Rodrigues, J.E.A.; Gil, A.M.; Barros, A.S.; Esteves, V.I. Analysis of non-aromatic organic acids by CE and direct mode with diode array detection. *Chromatographia*, **2009**, 70, 1737-1742.

Rodrigues, J.E.; Erny, G.L.; Barros, A.S.; Esteves, V.I.; Brandão, T.; Ferreira, A.A.; Cabrita, E.; Gil, A.M. Quantification of organic acids in beer by nuclear magnetic resonance (NMR)-based methods. *Analytica Chimica Acta*, **2010**, 674, 166-175.

Rodrigues, J.A.; Barros, A.S.; Carvalho, B.; Brandão, T.; Gil, A.M.; Silva Ferreira, A.C. Evaluation of beer deterioration by gas chromatography-mass spectrometry/ multivariate analysis: A rapid tool for assessing beer composition. *Journal of Chromatography A*, **2011**, 1218, 990-996.

Rodrigues, J.A.; Barros, A.S.; Carvalho, B.; Brandão, T.; Gil, A.M. Probing beer aging chemistry by nuclear magnetic resonance (NMR) and multivariate analysis (MVA). *Analytica Chimica Acta*, **2011**, 702, 178-187.

Book Chapters

Gil, A.M.; Rodrigues, J. High Resolution NMR tools for the analysis of beer and wine, in *Magnetic Resonance in Food Science – From Molecules to Man*, Eds. I.F. Farhat, P.S. Belton, G.A. Webb, The Royal Society of Chemistry, UK, Cambridge, 2007.

Gil, A.M.; Rodrigues, J. Methods for the Nuclear Magnetic Resonance spectroscopy characterization of beer, in *Beer in Health and Disease Prevention*, Ed. Victor Preedy, King's College, University of London, UK, Elsevier, 2008.

Rodrigues, J.E., Gil, A.M., NMR methods for beer characterization and quality control, in *Magnetic Resonance in Chemistry – Dealing in Complex Systems*, Eds. F. Capozzi, P.S. Belton, 2011 – accepted.

INDEX

Abbreviations and Symbols

5-HMF	5-hydroxymethylfurfural
ATR	attenuated total reflectance
BGE	background electrolyte
CE	capillary electrophoresis
CI	chemical ionization
COSY	correlation analysis spectroscopy
COW	correlation optimized warping
CTAB	cethyltrimethylammonium bromide
DDMP	2,3-dihydro-3,5-dihydroxy-6-methyl-4(H)-pyran-4-one
DTW	dynamic time warping
ECD	electron capture detection
EDTA	ethylenediaminetetraacetic acid
EI	electron ionization
EOF	electroosmotic flow
EPR	electron paramagnetic resonance
ERETIC	Electronic REference To access In vivo Concentrations
ESR	electron spin resonance
FFT	fast Fourier transform
FID	free induction decay
FT	Fourier transform
FT-IR	Fourier transform-infrared
GABA	γ-aminobutyric acid
GC	gas chromatography
HPLC	high performance liquid chromatography
HSQC	heteronuclear single quantum correlation
INEPT	insensitive nuclei enhanced by polarization transfer
IR	infrared
IS	internal standard
LC	liquid chromatography
LV	latent variable
LTP	lipid transfer protein
MBT	3-methyl-2-butene-thiol
MCCV	Monte Carlo cross-validation
MIR	mid-infrared
MLEV-17	Malcolm Levitt´s composite-pulse decoupling sequence

MS	mass spectrometry
MVA	multivariate analysis
NAD	nicotinamide adenine dinucleotide
NIPALS	non-linear iterative partial least squares
NIR	near-infrared
NMR	nuclear magnetic resonance
OP	outer-product
OPA	outer product analysis
PC	principal component
PCA	principal component analysis
PCT	principal component transform
PDC	2,6-pyridinedicarboxylic acid
PDVF	polyvinylidene fluoride
PLS	partial least squares
PLS-DA	partial least squares-discriminant analysis
ppm	parts per million
ppt	parts per trillion
PRESS	prediction error sum of squares
PTW	parametric time warping
rf	radio frequency
RI	Kovats retention index
RMSECV	root mean squares error of cross validation
RMSEP	root mean squares error of prediction
ROS	reactive oxygen species
R.T.	retention time
SNIF	site-specific natural ratio isotope fractionation
SPE	solid phase extraction
TOCSY	total correlation spectroscopy
TPPI	time-proportional receiver phase incrementation
TSP	3-(trimethylsilyl)propionate
TTAB	tetradecyltrimethylammonium bromide
TTAOH	tetradecyltrimethylammonium hydroxide
UPLC	ultra performance liquid chromatography
UV	ultraviolet
VDK	vicinal diketone
VIP	variable importance on the projection

1. THE BREWING PROCESS AND THE CHEMICAL COMPOSITION OF BEER

1.1 A brief history of beer

Beer is defined as a fermented beverage made from malted grains (usually barley), hops, yeast and water (Hughes and Baxter, 2001). In addition to malt from barley, other starch- and/or sugar-containing raw materials may be used, such as wheat malt, unmalted cereals (adjuncts), starch flour and fermentable sugars. The etymological origin of the word beer is thought to derive from *beor*, a West Germanic monastic borrowing of Vulgar Latin *biber "a drink, beverage"*, from Latin *bibere, "to drink"* (Klein, 1971; Boulton and Quain, 2001).

The art of brewing is as old as civilization, the oldest proven evidences of brewing dating from about 6000 BC and referring to Sumerian civilization. In fact, archeological studies performed in Mesopotamian region found documentation containing reference to beer production, such as an ancient hymn to the Sumerian goddess of brewing *Ninkasi*, which was also a beer recipe (Hornsey, 2003; Black *et al.*, 2004; Leick, 2010). The interest in beer expanded, along the centuries, mainly throughout the Middle East. For example, the Egyptians regarded beer not just as a drink, but also as product with religious and medical purposes (Bamforth, 2003; Hornsey, 2003) and it is thought that the Egyptians taught the brewing art to the Greeks and to the Romans. Although considered less important than wine, beer production evolved during the period of Roman domain, being mainly brewed in the outer areas of the Roman Empire by Saxons, Celtics, and Nordic and Germanic tribes. During the Middle Ages, the religious orders (i.e. Christian monks) preserved the brewing process as a craft and achieved several advances in the production process, including better storage techniques and the use of hops (Hornsey, 2003). It was only in the 12th century that the interest for brewing industry developed beyond monasteries walls, with the church also conceding the rights of brewing to private breweries. From the 16th century onwards, brewing expanded throughout the globe, becoming a global beverage.

The Industrial Revolution (18th century) had then a major impact on the brewing industry, allowing a tremendous increase in beer production and process efficiency (e.g. through mechanization) to be achieved as reviewed in (Hornsey, 2003; Priest and Stewart, 2006). Several scientific advances, such as the thermometer, saccharometer and artificial refrigeration, enabled improvements in process control and a greater knowledge of the final product to be achieved. Furthermore, in 1865, the notion of food preservation by pasteurization methods was established by Louis Pasteur. In 1881, Emil

Hansen was the first brewery scientist to isolate and classify specific brewery yeasts into bottom-fermenting lager strains (*Saccharomyces uvarum*, also called *Saccharomyces carlsbergensis*) and top-fermenting ale strains (*Saccharomyces cerevisiae*), devising methods for growing these yeasts in cultures free of other yeast and bacteria. This pure-culture technology was taken up quickly by lager brewers and settled the basis for beer production as we know it nowadays. In the 21st century, brewing industry has become a huge worldwide business, where the consumer has hundreds of beer types to choose, with brewers main concerns being focused on the beer quality control.

1.2 The brewing process

The brewing process involves several steps, as represented in Figure 1.1 (Hughes and Baxter, 2001). The first step, named malting, comprises three main operations: steeping, germination and kilning. The main goals of this process are the hydrolysis of the barley's cell walls and a portion of the endosperm protein, the increase of enzyme levels and the development of desirable malt color and flavor (Briggs *et al.*, 1981; Woonton *et al.*, 2005; Priest and Stewart, 2006). The malting process is initiated by steeping the grains in water, raising their water content from 12% to around 42-45 % by an osmotic process. This stimulates respiration in the embryo and hydrates the stores of starch in the endosperm of the barley grains, initiating the germination step. As the embryo activity increases, gibberellins (natural plant hormones) are produced and released, activating the aleurone layer of the grain to produce several hydrolytic enzymes, such as amylolytic, proteolytic and cellulytic enzymes (Boulton and Quain, 2001; Hughes and Baxter, 2001). These enzymes diffuse into the endosperm, beginning the process of breaking down the cell walls and the stores of protein, starch and lipids, providing nutrients for the new plant. At this stage, the barley grains are named green malt and the final operation (kilning) involves its drying, diminishing the water content of malt to approximately 4% (Boulton and Quain, 2001). This process stops germination and produces a stable form of malt, which can be safely stored in silos until needed for the brewing process. The temperature increase during kilning must be gradual so that the enzymes present in malt are not damaged. Simultaneously, aromatic and colored compounds are formed due to Maillard reactions (occurring on the heated malt) (Narziss, 2005; Bobalova *et al.*, 2010).

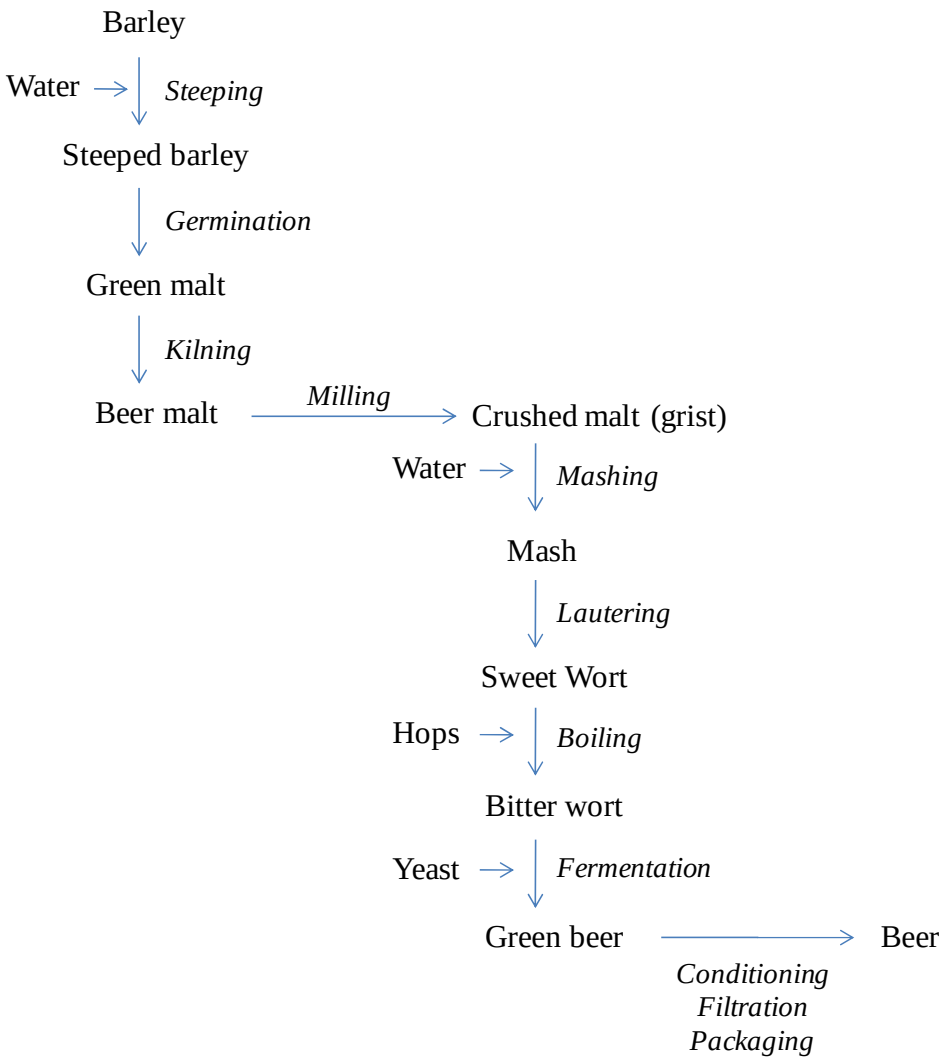

Figure 1.1. Simplified flow diagram of the brewing process (adapted from Hughes and Baxter, 2001).

In the brewery, the malt is crushed in a mill (milling step), releasing the content of the malt endosperm and reducing particle size, providing the required particle dimension for an efficient mashing stage (Boulton and Quain, 2001; Priest and Stewart, 2006). The resulting crushed malt, named *grist*, is mixed with hot water in the mash tun, at around 65ºC (mashing step). At this temperature malt starch gelatinizes, becoming more susceptible to enzyme attack (Hughes and Baxter, 2001). During mashing, the amylolytic enzymes, α- and β-amylases, break down the malt starch into fermentable sugars. Cereal starch consists of ~ 75 % amylopectin (branched glucose polymer, Figure 1.2a) and 25 % amylose (linear glucose polymer, Figure 1.2b). Both α- and β-amylases can hydrolyze α-1,4 bonds. The α-amylases attack the lengths of α-1,4 chains between branch points, releasing smaller, branched dextrins with long straight side-chains,

providing further substrate for β- amylase activity (Uchida *et al.*, 1991; Hughes and Baxter, 2001; Priest and Stewart, 2006).The β- amylases attack from the outer reducing ends of amylopectin and amylose molecules, releasing two glucose units (maltose) but stopping on the branched α-1,6 bonds of amylopectin.

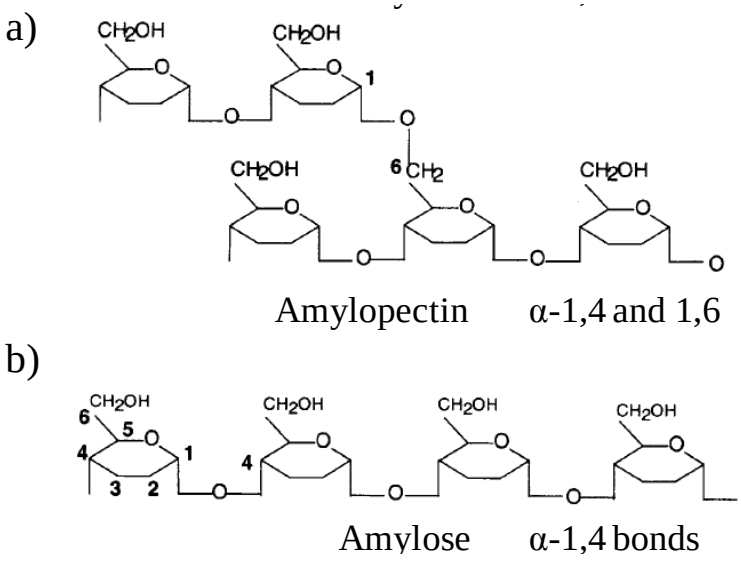

Figure 1.2. Molecular structures of amylose and amylopectin (adapted from Hughes and Baxter, 2001).

As a result of amylases activities, starch is reduced and a sweet syrupy liquid, called sweet wort, is obtained composed mainly by fermentable sugars (such as glucose, maltose, maltotriose and sucrose) and small branched dextrins (non fermentable). Protein hydrolysis also occurs in the mashing step, although in smaller extent when compared with starch degradation. Depending on the degree of protein modification (degree of breakdown of the protein-starch matrix), the enzyme content and the mashing conditions, peptides, amino acids and proteinaceous substances of high molecular weight are formed (Narziss, 2005). The sweet wort is then separated from the insoluble remains of the spent grains, in a lauter tun (lautering step). The resulting wort is delivered to a kettle where is subjected to vigorous boiling and hops or hop extracts are added (boiling step). The process has several functions, namely evaporation of excessive water in order to achieve the desired wort concentration (original extract), ensuring its sterilization and the inactivation of malt enzymes, coagulation of the proteins with formation of insoluble precipitates (called trub), and extraction of α-acids from hops into the wort and their isomerization into iso-α-acids (Boulton and Quain, 2001; Hughes and Baxter, 2001). In addition, the boiling process can result in the loss

of volatile components, such as essential oils from hops (by evaporation) and other undesirable volatiles derived from malt (for example, the decomposition of S-methylmethionine into dimethyl sulfide). Due to the elevated temperatures employed, Maillard reactions between the reducing sugars and the primary amines, particularly amino acids, are promoted with the formation of melanoidins, contributing to wort color (Hughes and Baxter, 2001).

Before fermentation, the trub and the hop material are removed from the wort (preventing wort-like, bitter or even harsh beer taste). In modern breweries, as most of the hops adopt the form of pellets or extracts, the whirlpool technique to clarify wort is mainly used. This technique consists in a cylindrical vessel into which the wort is pumped tangentially, causing the dump of the trub in the whirlpool base and enabling its separation from the wort. The resulting bitter wort is then cooled down to fermentation temperature using a heat exchanger (Boulton and Quain, 2001; Hughes and Baxter, 2001). Wort aeration and control of oxygen levels, particularly during the early stages of fermentation, are extremely important as they affect greatly yeast growth and, consequently, beer flavor (Priest and Stewart, 2006).

Fermentation is initiated by yeast addition to the wort (pitching). There are two main types of fermenting yeast: top-fermenting and bottom-fermenting. The top-fermenting yeast, *Saccharomyces cerevisiae*, is used in the production of ale beers, for which fermentation is carried out at temperatures around 16-25ºC, whereas bottom-form, *Saccharomyces carlsbergensis*, is used in the production of lager beer, which is fermented at around 7-14ºC and for longer periods of time. Generally, in ales production the yeast migrates to the surface of the beer during fermentation (top-fermenting), whereas in lagers the yeast settles to the base of the fermentation vessel (bottom-fermenting). However, nowadays, the most common fermentation systems for both ales and lagers beers production are the cylindro-conical vessels, where no distinction in the yeast settlements is seen (Bamforth, 2003). Yeast has the ability to adjust its metabolism to aerobic as well as anaerobic conditions, although the anaerobic reactions in beer occur in much higher extent (Narziss, 2005). Figure 1.3 shows a simplified scheme of some redox-balancing reactions occurring in yeast metabolism.

During beer fermentation, yeast takes up amino acids (used for cell growth) and sugars (mainly metabolized to produce energy with the formation of ethanol and carbon dioxide, under anaerobic conditions) from wort (Hughes and Baxter, 2001). Yeast

metabolism and growth also originates the production of several flavor-active compounds (by-products), mainly higher aliphatic and aromatic alcohols, esters, organic acids, carbonyl compounds, sulphur-containing compounds and polyhydric alcohols, all of which important for the properties and quality of the resulting beer (Narziss, 2005). Thus, yeast is responsible for much of the unique character which distinguishes one beer from another (Hughes and Baxter, 2001). Once the yeast has fermented all the available sugars, its metabolism slows down and yeast cells start to settle in the vessel. When the fermentation ceases, the vessel is cooled down to around freezing, causing yeast cells to flocculate, together with the coagulation of proteins. The bulk of the yeast is then separated from the fresh beer in process known as racking. This freshly produced beer, referred to as green beer, still contains undesirable flavor-active compounds which are removed by conditioning (Priest and Stewart, 2006).

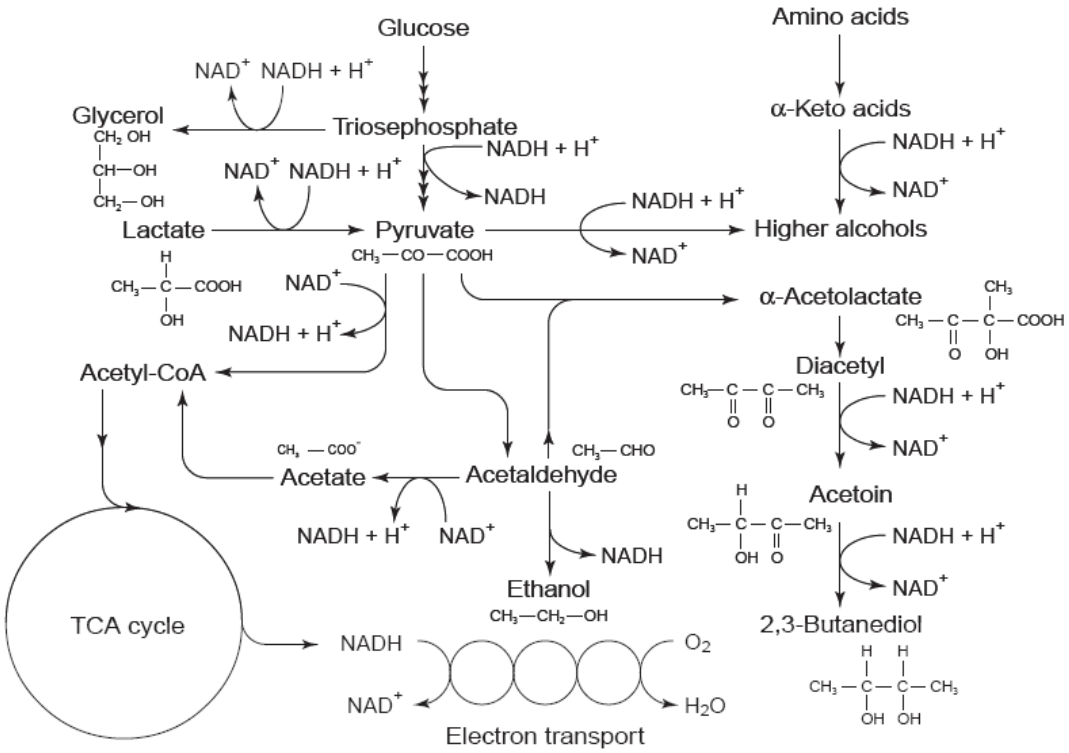

Figure 1.3. Simplified scheme showing redox balancing reactions in yeast metabolism (reproduced from Briggs, 2004).

Conditioning comprises three main steps: maturation, stabilization and clarification. In the maturation step, a secondary fermentation occurs, with the remaining yeast using fermentable sugars (residues and/or by adding sugar) to produce carbon dioxide, enabling the carbonating of beer and purging out unwanted volatile

compounds. Moreover, yeast can chemically remove certain undesired flavor-active components, such as the vicinal diketones (VDK´s) diacetyl and 2,3-pentanedione, reducing them into compounds with no relevant flavor attributes. Beer stabilization and clarification is obtained by storing the product at 0-2 ºC. This stabilization will originate the flocculation of the remaining yeast, removed by centrifugation or by adding fining agents, and the precipitation and subsequent removal of polypeptide-polyphenols complexes, using chillproofing agents, enabling improving beer colloidal stability (Goldammer, 2008).

Beer is then filtered and possible adjustments of beer characteristics, such as CO_2 levels, color, bitterness and alcohol strength, can be performed before the final step of packaging (Priest and Stewart, 2006). The latter is usually done into either can, bottle (glass or plastic), keg, cask or bulk tanks (Bamforth, 2003). This process is subject to various legal regulations of quality control and has to be extremely efficient in order to avoid any entrance of oxygen in the system as it would originate beer staling.

Different combinations of ingredients (raw materials) and production processes give rise to an extensive variety of beers, with ales, stouts and lagers being defined as the main types of beer (the first two deriving from top fermentation and the latter from bottom fermentation). In the last years, the brewing industry quest for new products lead to the development of new beer types differing in some specific characteristics such as ice, dry, light, draft and low- or non-alcohol beers (Bamforth, 2003; Priest and Stewart, 2006).

1.3 Beer quality: main characteristics and chemical composition

The different combinations of ingredients, brewing processes and brewery design give rise to a chemically complex product consisting in hundreds of components present in a wide range of concentrations, the most abundant being water, ethanol and carbohydrates (Hughes and Baxter, 2001; Bamforth 2009). The main constituents and concentration levels of a typical beer are listed in Table 1.1. Beer alcohol content ranges from < 0.05%, in alcohol free beers, to 12.5% in very strong beers. Ethanol is present in most beers at levels, at least, two orders of magnitude higher than other alcohols, contributing importantly to flavor. Glycerol is the second most abundant alcohol, with concentration ranging from 12–20 g/100 mL and other alcohols, such as propanol, isobutanol, 2- and 3-methylbutanol, 2-phenylethanol and tyrosol, are also present in

beer, although in smaller quantities, being secondary metabolites of yeast amino acids metabolism (Hughes and Baxter, 2001; Bamforth, 2009). These alcohols also affect beer flavor and can act as precursors of flavor-active esters.

Table 1.1. Typical beer composition (adapted from Hughes and Baxter, 2001).

Components	**Typical levels (g/ 100 mL)**
Water	92-95
Alcohol	2.5-3.5
Total carbohydrates	1.5-3.0
Residual sugars	< 0.2
Non-starch polysaccharides	0.3-1.0
Protein-derived components	0.2-0.6
Lipids	Negligible
Vitamins and other micronutrients	0.002
Polyphenols and hop components	0.002-0.06
Organic acids	< 0.15
Minor components (flavor-active)	
Esters	
Aldehydes	
Vicinal diketones (VDK's)	
Sulfur compounds	

Carbohydrates are the major non-volatile components of beer, with typical amounts of 1.5-3.0 g/100 mL. Their source is starch-rich cereals, mainly malted barley, although other cereals like wheat, rice, maize, oats and sorghum are used in other beers (Hughes and Baxter, 2001). Carbohydrates comprise mainly fermentable sugars (e.g. sucrose, glucose, fructose, maltose, maltotriose and trehalose) and dextrins (i.e. unfermentable oligomers with varying degrees of polymerization and branching patterns) (Uchida *et al.*, 1991; Hughes and Baxter, 2001). The residual sugars, i.e. the fermentable sugars, arise either from the malt (although the majority are consumed in the fermentation step) or from their addition in the conditioning step to stimulate a secondary fermentation (to bring the beer up to specified carbon dioxide levels). Lager beers are generally more fully fermented than ales, thus containing lower amounts of residual sugars, which contribute directly to beer sweetness. Oligosaccharides with more than four glycosyl units induce little sweetness, but are important to beer body and mouthfeel. The non starch-polysaccharides in beer (present in the range of concentration

between 0.3-1.0 g/100 mL) arise from the β-glucan cell walls in barley. Although the high molecular weight β-glucans are degraded during malting, the smaller and soluble glucans remaining in beer contribute do dietary fibre. Beer is also a good source of vitamins, with niacin, panthothenic acid, pyridoxine and riboflavin being the most abundant (Hughes and Baxter, 2001).

Beer typically contains between 0.2-0.6 g/100 mL of proteins, peptides and amino acids, which mainly originate from the malted barley. Many of the larger proteins are precipitated and removed during wort boiling process, while most of the free amino compounds are taken up by the yeast. Thus, most of the proteins components are in the form of peptides and polypeptides, which are readily digested. Nevertheless, most beers contain all the essential amino acids, at levels generally varying between 5-10 mg/mL (Hughes and Baxter, 2001).

Lipids are a significant component of malt, as they represent typically around 3% (w/w) of dry barley grain (Palmer, 2006), mainly in the form of triglycerides. The limited lipase activity during malting and mashing releases some fatty acids, mainly unsaturated linoleic and linolenic acids, into the wort. Most of these are taken up for yeast growth, so that the main lipids left in beer are medium linear fatty acids caproic, caprylic and capric acids, contributing with a rancid or goaty flavor for the organoleptic characteristics of beer (Briggs, 2004; Horak *et al.*, 2009).

In relation to phenolic composition, beer contains a wide range of compounds, such as flavonoids, flavonols, anthocyanidins, proanthocyanidins, and phenolic acids and alcohols, derived both from hops and malt. These compounds play an important role on the stability and flavor of beer, contributing directly to flavor, astringency, haze body and fullness (Briggs, 2004; Callemien and Collin, 2009). Some phenolic structures can also impart antioxidant activity with potential beneficial to health.

Hop acids, consisting of α- and β-acids (alicyclic phenolic acids) are of great importance in the organoleptic characteristics of beer. Indeed, the isomerization products of hop α-acids during wort boiling results in the formation of iso-α-acids (Figure 1.4), which are mainly responsible for beer bitterness and foam properties (Verzele and De Keukeleire, 1991; De Keukeleire, 2000; Nord *et al.*, 2003). The resulting iso-α-acids are, in fact, a mixture of closely-related compounds, three pairs of cis- and trans- isomers, each pair deriving from a single α-acid. Hops also contain β-acids, which are considered as minor contributors to bitterness.

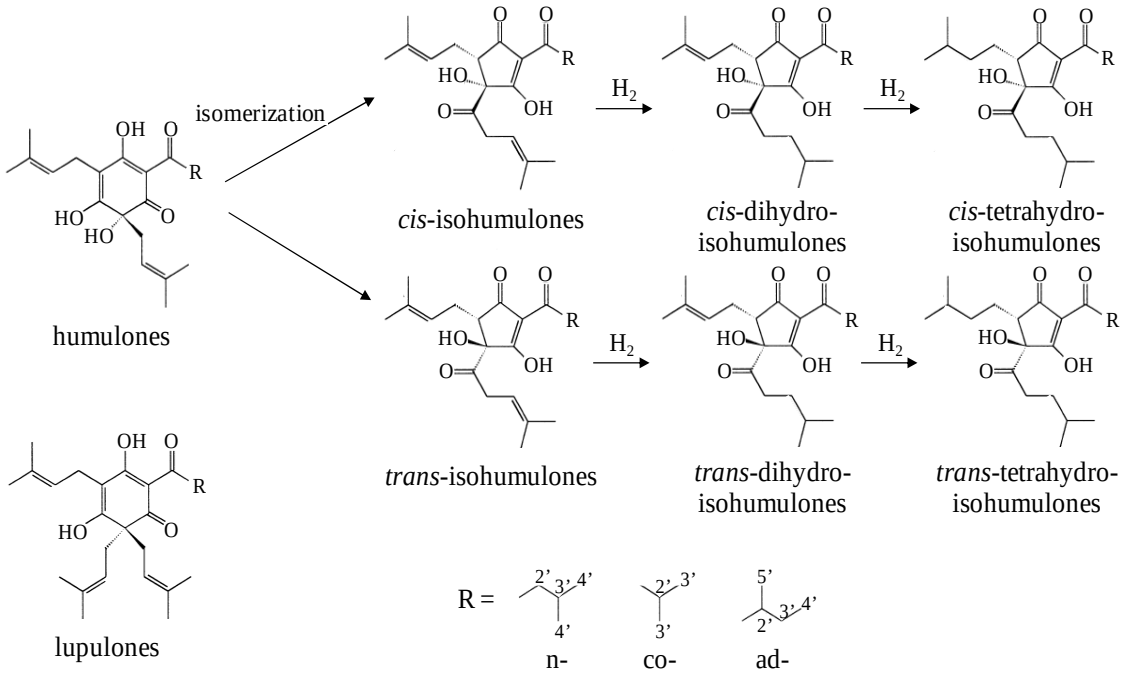

Figure 1.4. Structure of the main hop bitter acids: lupulones (β-acids), humulones (α-acids), iso-humulones (*cis*- and *trans*- configuration) and their derivatives (hydrogenated isohumulones) (reproduced from Pusecker *et al.*, 1999).

The presence of organic acids in beer has a sourness impact in flavor. The major acids found in beer include acetic, citric, lactic, malic pyruvic and succinic acids, with amounts varying between 10 and 300 mgL^{-1} The levels of acids present in beer depend on the vigor of fermentation, with faster fermentations originating higher amounts of these acids.

Other components, although present in low amounts in beer, play important roles in the organoleptic characteristics of beer. Fermentation is the decisive step to the final beer flavor expression, with the formation of a whole array of volatile compounds with organoleptic relevance, the profile of which are a function of i) the yeast strains used and ii) the wort boiled composition that is pitched. The most relevant compounds formed are the volatiles esters, alcohols, aldehydes, vicinal diketones (VDK's) and sulfur compounds, as well as organic acids (see Table 1.2).

Esters are the most important group of fermentation-derived flavor-active components and result from the esterification of an organic acid with an alcohol. They impart fruity, floral and solvent-like flavors and aroma to beers (Table 1.2). The most abundant ones are those derived of organic acids with ethanol (e.g. ethyl acetate, ethyl hexanoate and ethyl octanoate), as expected since ethanol is the most abundant alcohol

in beer (Boulton and Quain, 2006; Casey, 2007). The levels of esters in beer are influenced by various factors, such as the gravity (density) of the wort and the amount of oxygen to which yeast is exposed. However, perhaps the major factor affecting the extent of ester produced is the yeast strain itself, with some strains more readily generating esters than others.

Regarding alcohols, as mentioned before, ethanol contributes directly for flavor, with a warming and alcoholic character, as well as plays a role in the flavor perception of other components (in partitioning of flavor components between the liquid beer, foam and the headspace). The flavor-active higher alcohols (e.g. propanol, isobutanol and 2- and 3-methylbutanol) are important as they impart a warming, alcoholic flavor and aroma, and can act as precursors of flavor-active esters (Hughes and Baxter, 2001). Yeast strains are the most important factor affecting the production, with ale strains producing more amounts higher alcohols than lager strains. Furthermore, conditions which favor increased yeast growth, such as aeration or oxygenation, will also promote the formation of higher alcohols.

The production of aldehydes during fermentation is closely related to the formation of higher alcohols. Aldehydes are much more flavor-actives than their corresponding alcohols, and are generally considered to contribute negatively for flavor attributes of beer (Bamforth, 2009). The major aldehyde in beer is acetaldehyde, which gives an emulsion paint or green apple aroma and taste to beer. Other aldehydes can be responsible for flavor defects in aged beer, either by higher alcohol oxidation, by Strecker degradation of amino acids or by Maillard reactions.

Similarly, the vicinal diketones (VDK's) in beer usually impart a negative impact to flavor. The major VDK in beer is 2,3-butanedione (diacetyl), but also significant quantities of 2,3-pentanedione are produced during fermentation. Diacetyl typically has a butterscotch flavor and is also associated with a sweetness perception. Elimination of VDKs from beer depends on fermentation management. During fermentation, extracellular VDK's can be metabolized by yeast cell reductases, being reduced sequentially first to acetoin and then to 2,3-butanediol (Hughes and Baxter, 2001). These reduced metabolites are much less flavor-active than the VDK's, being more tolerated in beer.

Table 1.2. Significant flavor-active alcohols, aldehydes, esters and vicinal diketones (VDK's) as well as acids found in beer (adapted from Hughes and Baxter, 2001).

	Typical levels (mgL^{-1})	Flavor threshold (mgL^{-1})	Flavor descriptors
		Esters	
2-phenylethyl acetate	0.05-2.0	3.0	Roses, honey, apple, sweet
Ethyl acetate	10.0-60.0	30.0	Solvent-like, sweet
Ethyl hexanoate	0.1-0.5	0.2	Apple, fruity, sweet
Ethyl nicotinate	1.0-1.5	2.0	Grainy, perfume
Ethyl octanoate	0.1-1.5	0.5	Apple, tropical fruit, sweet
Isoamyl acetate	0.5-5.0	1.0	Banana, ester, solvent
		Alcohols	
1-octen3-ol	0.03	0.2	Fresh-cut grass, perfume
n-propanol	3-16	700	Alcoholic
2-decanol	0.005	0.015	Coconut, aniseed
2-methylbutanol	8-30	65	Alcoholic, vinous, banana
2-propanol	3-6	1500	Alcoholic
2-phenylethanol	8-35	125	Roses, bitter, perfumed
3-methylbutanol	30-70	70	Alcoholic, vinous, banana
Ethanol	20000-80000	14000	Alcoholic, strong
Glycerol	1200-2000	-	Sweetish, viscous
Methanol	0.5-3.0	10000	Alcoholic, solvent
Tyrosol	3-40	200	Bitter, chemical
		Aldehydes	
Acetaldehyde	2-20	25	Green, paint
Trans-2-butenal	0.003-0.02	8	Apple, almond
2-methylpropanal	0.02-0.5	1	Banana, melon
C5 aldehydes	0.01-0.3	ca. 1	Grass, apple, cheese
Trans-2-nonenal	0.00001-0.002	0.0001	Cardboard
Furfural	0.01-1	200	Papery, husky
5-methylfurfural	< 0.01	17	Spicy
5-HMF	0.1-20	1000	Aldehyde, stale
		VDK's	
2,3-butanedione	0.01-0.4	0.07-0.15	Butterscotch
3-hydroxy-2-butanone	1-10	17	Fruity, mouldy, woody
2,3-butanediol	50-150	45000	Rubber, sweet, warming
2,3-pentanedione	0.01-0.15	0.9	Butterscotch
3-hydroxy-2-pentanone	0.05-0.07	-	-
		Organic acids	
Acetic	30-300	175	Acid, vinegar
Butanoic	0.5-1.5	2.2	Butter, cheesy, sweat
Citric[1]	56-230	-	-
Lactic	20-80	400	Acid
Malic[1]	14-105	-	-
2-methylbutanoic	0.1-0.5	2	Cheese, old hops, sweat
3-methylbutanoic	0.1-2	1.5	-
2-methylpropanoic	0.1-2	30	Sweat, bitter, sour
Octanoic	2-12	15	Caprylic, goaty
Pentanoic	0.03-0.1	8	Sweat, body odor
Propanoic	1-5	150	Acid, vinegar, milk
Pyruvic	15-150	300	Acid, salt, forage
Succinic	16-140	-	-

[1] Concentration values extracted from Montanari *et al.*, 1999.

Some of the most easily recognized flavors in beer are sulfur-containing compounds, such as thiols, sulfides, disulfides, thioesters and sulfur heterocyclic compounds (Bamforth, 2009). These compounds differ with the type of fermentation, with some ale beers having considerable amounts of hydrogen sulfide giving a rotten eggs or drains attributes to beer flavor, whereas lager beers tends to have a more complex sulfur character. The major sulfur-flavor compound is usually dimethyl sulfide, inducing a sweetcorn flavor. Other sulfides are extremely relevant, such as methional and methionol, characterized by flavor reminiscent of raw and cooked potato, respectively. Thiols are highly flavor active, being the most unpleasant one the 3-methyl-2-butene-thiol (MBT). This compound is flavor-active at low part per trillion (ppt) levels, being formed by light-induced degradation of hop-derived iso-α-acids. Thioesters, especially methyl and ethyl thioacetates, are less flavor-active and have less offensive flavor attributes, such as cooked vegetable. Finally, the heterocyclic sulfur compounds are generally less well-characterized, however, flavor attributes such as roasted and meaty are common for such compounds.

1.4 Beer stability

Beer shelf-life is mostly determined by biological (contamination by bacteria, yeast and mycelia fungi) and physical/chemical stability (divided into physical or haze formation, and foam, light and flavor chemical processes) (Priest and Stewart, 2006). In this thesis, the studies related with the beer aging process were focused on the factors affecting the flavor stability and, therefore, this issue will be described with more attention.

Haze formation

In haze formation, the primary reaction is the polymerization of polyphenols and their bindings with specific proteins. The balance between polyphenols and the specific proteins will dictate the physical stability, with haze forming slowly or no form when either of these components is removed. Haze formation is highly influence by storage temperature, the presence of oxygen, heavy metal ions and light (Priest and Stewart, 2006; Bamforth, 2009). Beer shelf-life can also be prolonged, avoiding haze formation, by using stabilizing agents.

Foam stability

Foam stability is widely considered as reflecting the quality of the product. There are many foam-promoting compounds that play an important role in foam formation and stability, such as iso-α-acids, hops, metal ions, protein Z, LTP1 and various hordein-derived species, and polysaccharides (Bamforth, 2009).

Light Stability

Light instability in beer is a consequence of hops acids. Indeed, when exposed to light, the iso-α-acids are decomposed by light-induced reactions producing MBT, which is a unpleasant flavor-active compound characterized by a skunky aroma (Irwin *et al.*, 1991; Priest and Stewart, 2006).

Flavor stability

The flavor stability of beer, considered as the biggest quality challenge for the brewing industry, is of paramount importance in order to fulfill the consumers expectations. Indeed, a variety of compositional changes may occur during storage, either by formation and/or degradation reactions, resulting in the loss of the expected beer flavor (Vanderhaegen *et al.*, 2006). The factors affecting beer aging, closely connected to flavor degradation, can be divided into intrinsic factors, such as oxygen content (Depraetere *et al.*, 2008; Savel *et al.*, 2010), pH (Gijs *et al.*, 2002a; François *et al.*, 2006), antioxidants contents (Lermusieau *et al.*, 2001; Gijs *et al.*, 2002b) and precursors concentrations of key odorants (Gijs *et al.*, 2002a; Zufall and Tyrell, 2008); and extrinsic factors, such as packaging (Kuchel *et al.*, 2006), temperature (Vanderhaegen *et al.*, 2003; Soares da Costa *et al.*, 2004) and light (Savel *et al.*, 2010). From these, oxygen content and temperature has been the most studied factors and are discussed in more detail following.

Oxidation degradation of beer is considered to be a major source of flavor instability, with the reactive oxygen species (ROS) having an important role in potentiating flavor deterioration. The oxygen conversion into the ROS is illustrated in Figure 1.5. Briefly, oxygen in the ground state ($^3O_2$) is quite stable and will not easily react with organic molecules. However, in the presence of agents, such as transition metal ions (Fe^{2+} and Cu^{+}), enzymes and light, it can capture an electron and form the superoxide anion (O_2^{-}). The superoxide can be protonated to form the perhydroxyl

radical ($OOH^{\bullet}$) which has much higher reactivity. Since the pKa of this reaction is 4.8, in beer (generally, beer pH values vary between 4.0 and 4.5) the majority of the superoxide will be in perhydroxyl form. The superoxide anion can also be reduced by Fe^{2+} or Cu^{+} to the peroxide anion (O_2^{2-}), which is readily protonated to hydrogen peroxide (H_2O_2). The hydroxyl radicals ($OH^{\bullet}$) can be produced from both the hydrogen peroxide and the superoxide anion by metal induced reactions, such as Fenton and the Haber-Weiss reaction (Vanderhaegen *et al.*, 2006; Bamforth, 2009). The reactivity of the ROS increases with their reduction status: O_2^{-} < $OOH^{\bullet}$ < $OH^{\bullet}$, with their concentration increase being related with higher Fe^{2+} and Cu^{+} content, oxygen concentration and storage temperatures (Kaneda *et al.*, 1992; Vanderhaegen *et al.*, 2006).

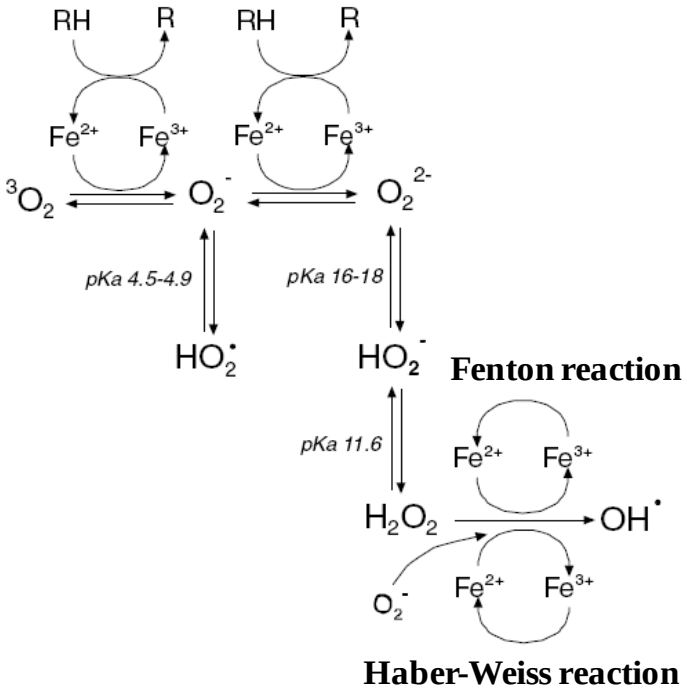

Figure 1.5. Reactions producing reactive oxygen species (ROS) in beer (reproduced from Kaneda *et al.*, 1999).

Apart from oxygen, the most dramatic impact on flavour stability is the temperature of storage. The relationship between temperature storage and a chemical reaction can be described as (Bamforth, 2009):

$$k_{t+10} = 2 \sim 3\ k_t \qquad \text{[Eq. 1.1]}$$

where k_t is the rate of a chemical reaction at a given temperature and k_{t+10} is the rate of that reaction when the temperature is raised 10ºC, indicating that if the temperature of beer is raised 10 ºC, the reactions rate is increased two to threefold. The relevant impact that temperature has on the development of stale attributes is dramatic, as shown in Figure 1.6. This relationship (Eq. 1.1) is widely employed on studies of beer aging, with forced aging protocols being based on thermal treatment used to mimic natural aging (Chevance *et al.*, 2002; Varmuza *et al.*, 2002; Guido *et al.*, 2004; Vanderhaegen *et al.*, 2004a; François *et al.*, 2006). The use of forced aging protocols is advantageous as it allows measuring the impact of technological variables on chemical mechanisms, enabling the mimicking of modifications in controlled laboratory conditions and obtaining relevant compositional information in a short period of time. Nevertheless, it is important to underline that thermal treatment of beer could originate a disproportionate development of staling compounds when compared with natural aging (Depraetere *et al.*, 2008).

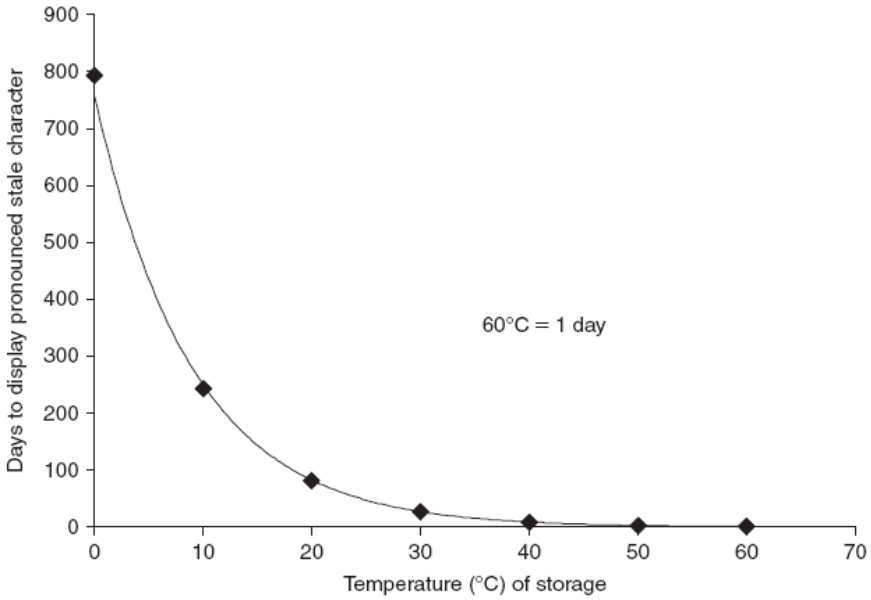

Figure 1.6. Impact of temperature on flavor deterioration of beer (reproduced from Bamforth, 2009).

The main compositional changes occurring in beer aging are well recognized. Table 1.3 lists the most relevant compounds varying with beer aging, both flavor-active associated with flavor deterioration as well as non-volatile compounds. The content variations of these known compounds are related with several different chemical mechanisms, such as the Maillard reactions, Strecker degradation of amino acids,

oxidation of fatty acids and higher alcohols, aldol condensation, hop bitter acids (iso-α-acids, α-acids and β-acids) and polyphenols degradation, acetalization of aldehydes, synthesis and hydrolysis of volatile esters and the formation of (E)-β-damascenone (Vanderhaegen *et al.*, 2006). The most relevant of these reactions are described below.

Table 1.3. Overview of the most relevant known volatile and non-volatile compounds varying during storage (reproduced from Vanderhaegen *et al.*, 2006).

Chemical class	**Compounds**
Linear aldehydes	• Acetaldehyde • (E) -2-octenal/ (E) -2-nonenal/ (E,E)-2,4-decadienal/ (E,E)-2,6-nonadienal/
Strecker aldehydes	• 2-methylbutanal/ 3-methylbutanal/ 2-methylpropanal/ phenylacetaldehyde/ benzaldehyde/ methional
Ketones	• (E)-β-damascenone • 3-methyl-2-butanone/ 4-methyl-2-butanone/ 4-methyl-2-pentanone • diacetyl/ 2,3-pentanedione
Cyclic acetals	• 2,4,5-trimethyl-1,3-dioxolane/ 2-isopropyl-4,5-dimethyl-1,3-dioxolane/ 2-isobutyl-4,5-dimethyl-1,3-dioxolane/ 2-sec butyl-4,5-dimethyl-1,3-dioxolane
Heterocyclic compounds	• Furfural/ 5-hydroxymethylfurfural (5-HMF)/ 5-methylfurfural/ 2-acetylfuran/ 2-acetyl-5-methylfuran/ 2-propionylfuran/ furan/ furfuryl alcohol • Furfuryl ethyl ether/ 2-ethoxymethyl-5-furfural/ 2-ethoxymethyl-2,5-dihydrofuran • Maltol • Dihydro-5,5-dimethyl-2(3H)-furanone/ 5,5-dimethyl-2(5H)-furanone • 2-acetylpyrazine/ 2-methoxypyrazine/ 2,6-dimethylpyrazine/ trimethylpyrazine/ tetramethylpyrazine.
Ethyl esters	• Ethyl-2-methylbutyrate/ ethyl-2-methylbutyrate/ ethyl-2-methylpropionate • Ethyl nicotinate/ diethyl succinate/ ethyl lactate/ ethyl phenylacetate/ ethyl formate/ ethyl cinnamate
Lactones	• γ-nonalactone/ γ-hexalactone
Sulfur compounds	• Trimethyl sulfide/ 3-methyl-3-mercaptobutylformate
Non volatile compounds	• Iso-α-acids/ phenolic compounds/ amino acids

Regarding Maillard reactions (reactions between reducing sugars and proteins, peptides, amino acids or amines) (Figure 1.7), several heterocyclic compounds are well known markers of beer aging, having been described as increasing with time at an approximately linear rate which varies logarithmically with the storage temperature (Madigan *et al.*, 1998). Quantitatively, 5-hydroxymethylfurfural (5-HMF) is one of the

most important heterocyclic staling compounds, being formed by combination of a hexose and an amino acid group. Other relevant aging heterocyclic compounds include furfural, furfuryl ethyl ether and furfuryl alcohol (Vanderhaegen *et al.*, 2003; Vanderhaegen *et al.*, 2004a; Vanderhaegen *et al.*, 2004b).

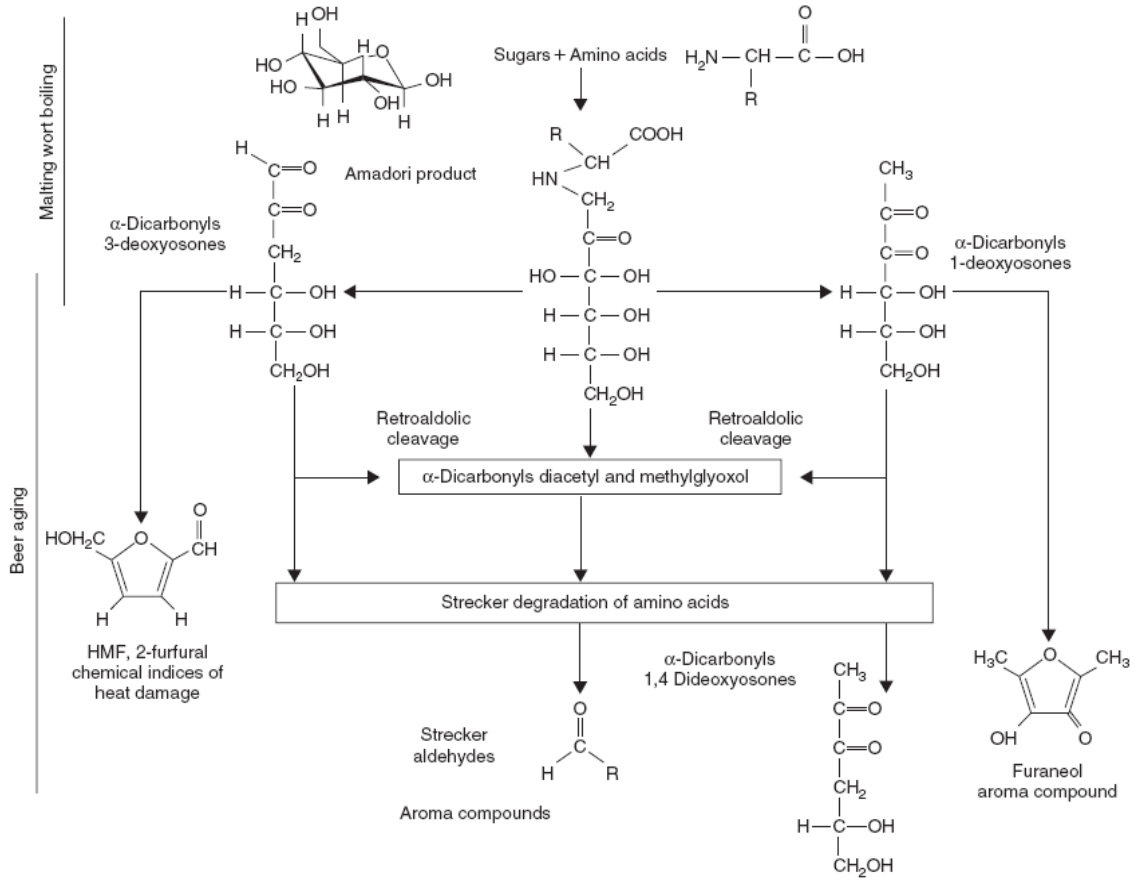

Figure 1.7. Overview of Maillard reaction in beer (reproduced from Bravo *et al.*, 2002).

Strecker degradation of amino acids, reaction between amino acids and α-dicarbonyl compounds, involves transamination, followed by decarboxilation of the subsequent α-ketoacid, resulting in a flavor-active aldehyde with one carbon atom less than the corresponding amino acid. The most flavor-active aldehydes formed during storage are methional (cooked potato aroma) and phenylacetaldehyde (honey-like aroma) from the degradation of respectively methionine and phenylalanine (Gijs *et al.*, 2002a; Soares da Costa *et al.*, 2004). Furthermore, an increase in the amounts 2-methylpropanal, 2-methylbutanal and 3-methylbutanal by degradation of valine, isoleucine and leucine, respectively, were also identified (Soares da Costa *et al.*, 2004), being responsible for a malty aroma. The Strecker aldehydes formation is greatly

affected by the temperature of storage and the levels of dissolved oxygen, especially when close to saturation (Silva Ferreira *et al.*, 2002).

The oxidation of fatty acids has received much attention, particularly due to the formation of (E)-2-nonenal. There are two known pathways to the oxidation of unsaturated fatty acids: enzymatic oxidation (occur during mashing) and autoxidation. Recent studies showed strong evidence that lipid oxidation does not occur after bottling at normal storage temperatures, nor can nonenol oxidation account for the presence of (E)-2-nonenal (Lermusieau *et al.*, 1999). In fact, (E)-2-nonenal content in aged beer originates from autoxidation during wort boiling (Noel *et al.*, 1999) and enzymatic action during mashing (Liegeois *et al.*, 2002).

The oxidation of higher alcohols, such as 2-methylpropanol (isobutanol), 2-methylbutanol (pentanol) and 3-methylbutanol (isopentanol), into the corresponding aldehydes valine, leucine and isoleucine occur during beer aging, being favored by the presence of high oxygen levels and temperature and/or low pH values (Hashimoto, 1972; Vanderhaegen *et al.*, 2006). Aldol condensation in beer is possible under the mild conditions existing in beer during storage (Hashimoto and Kuroiwa, 1975), with this reaction being catalyzed by amino acids through the formation of an intermediate amine group.

Hop bitter acids (α-acids, β-acids and iso-α-acids) can undergo go oxidative degradation resulting in the formation of alkanones, alkenals and alkadienals of varying length and structure (Hashimoto and Eshima, 1979). This degradation not only affects beer bitterness and astringency characteristics, but also results in the formation of by-products, some of them known flavour-active compounds such as the 3-methyl-2-butene-1-thiol (Irwin *et al.*, 1991).

Polyphenols, in beer, easily react with ROS and it is believed that simple polyphenols polymerize to high molecular weight species (Gardner and McGuinness, 1977). During beer storage, interactions between phenolic polymers and proteins can also occur, resulting in the formation of insoluble complexes and haze (Vanderhaegen *et al.*, 2006).

The cyclic acetals, 2,4,5-trimethyl-1,3-dioxolane, 1-isopropyl-4,5-dimethyl-1,3-dioxolane, 2-isobutyl-4,5-dimethyl-1,3-dioxolane and 2-sec butyl-4,5-dimethyl-1,3-dioxolane, originate from the acetalization of aldehydes, i.e. the condensation reaction between 2,3-butanediol and an aldehyde, acetaldehyde, isobutanal, 3-methylbutanal and

2-methylbutanal, respectively (Peppard and Halsey, 1982). In beer, an equilibrium between 2,4,5-trimethyl-1,3-dioxolane, acetaldehyde and 2,3-butanediol is reached quite rapidly, with the increase of acetaldehyde during aging causing the 2,4,5-trimethyl-1,3-dioxolane concentration to increase very similarly (Vanderhaegen *et al.*, 2003).

Esters are known to either increase, through hydrolysis, or decrease, by esterification, their concentration with the beer aging process. In fact, the acetate esters of higher alcohols, e.g. isoamyl acetate and ethyl acetate, were shown to decrease due to hydrolysis (Vanderhaegen *et al.*, 2007). On the other hand, the esterification of ethanol with organic acids also occurs, with increased concentration of diethyl succinate, ethyl pyruvate, ethyl lactate, ethyl nicotinate and ethyl phenylacetate identified in aged beers (Vanderhaegen *et al.*, 2003, Vanderhaegen *et al.*, 2007).

Finally, (E)-β-damascenone, a carotenoid-derived carbonyl compound with a fruit-flowery flavor, also increases with beer aging. The hydrolysis of glycosides and the conversion from allene triols and dihydroxyacetylene have been suggested to explain the formation of (E)-β-damascenone during aging (Chevance *et al.*, 2002; Gijs *et al.*, 2002a).

1.5 Beer analysis

1.5.1 Commonly used methods

Beer analysis is important for the evaluation of i) organoleptic characteristics, ii) general quality, iii) nutritional aspects, and iv) safety of the final product. Typically, beer quality control in brewing industry is assessed through sensory evaluation of the organoleptic characteristics such as color, taste, appearance, flavor and aroma, with some analytical measurements, such as through photometry, enzymatic analyses and chromatographic methods being also used. The sensory analysis of beer has been a major focus of attention to brewing researchers, particularly concerning the beer stability issue. The first studies reporting the use of sensory response to investigate the chemical aspects of beer staling has been described in the 1970s (for example: Drost *et al.*, 1971; Meilgaard, 1972; Dalgliesh, 1977). Throughout the years, sensory analysis has enabled a better understanding of the effects of aging in the organoleptic properties of beer. Sensory response data have been mainly employed in assessing the factors affecting the flavor degradation rate, such as the oxygen levels and beer pH (Guyot-Declerck *et al.*, 2005; Depraetere *et al.*, 2008; Malfliet *et al.*, 2009), improvements of

the flavor thresholds values (Saison *et al.*, 2009) and in predictive studies of organoleptic stability (Daems and Delvaux, 1997; Guido *et al.*, 2007; Liu *et al.*, 2008).

The urge for objective analytical methods for on-line assessment of beer quality parameters, led to an increasing use of analytical methods for the characterization of the chemical composition of beer, following two different strategies: targeted and untargeted approaches. In a targeted approach, the aim is to, simultaneously, detect and accurately determine relative abundancies or concentrations of specific, pre-determined set of known compounds for assessment of their behavior, under determined conditions. On the other hand, untargeted approaches are used for global metabolic analysis, i.e. for comprehensive analysis of all the measurable analytes in a sample (including analyte identification of unknown signals) to obtain patterns or fingerprints related with the studied phenomena without, necessarily, identifying nor quantifying specific compounds (Cevallos-Cevallos *et al.*, 2009). Untargeted studies are usually coupled to chemometric methods to compress the data into a small set of signals, enabling the use of a larger number of data, in a high throughput manner, being more successful in identifying metabolites that play a key role in specific biochemical processes, flavor characteristics or quality/authenticity studies.

Until recently, the chromatographic methods of high-performance liquid chromatography (HPLC) and gas chromatography (GC) were the main tools in beer analysis enabling the determination of several compounds in beer, such as carbohydrates, alcohols, amino acids, phenolic compounds, carboxylic acids, sulfur compounds and bitter acids. Chromatographic methods shown a wide range of applications in beer analysis, such as beer aging studies (Chevance *et al.*, 2002; Soares da Costa *et al.*, 2004; Guido *et al.*, 2004; Vanderhaegen *et al.*, 2004a) and the compositional profiling of the volatile (Pinho *et al.*, 2006; Silva *et al.*, 2008) and non-volatile components (Vinogradov and Bock, 1998; Erbe and Bruckner, 2000; Nord *et al.*, 2003; Kabelova *et al.*, 2008).

In recent years, other analytical techniques have gained their space in beer analysis, such as capillary electrophoresis (CE), mainly used in quantification studies (Klampfl, 1999; Cortacero-Ramirez *et al.*, 2005; Erny *et al.*, 2009), nuclear magnetic resonance (NMR), applied in the chemical characterization (Vinogradov and Bock, 1998; Gil *et al.*, 2003) and quality control studies (Duarte *et al.*, 2004; Nord *et al.*, 2004; Lachenmeier *et al.*, 2005), infrared (IR) spectroscopy applied to quality control studies

(Duarte *et al*., 2004; Inon *et al*., 2006; Lachenmeier, 2007), and electron paramagnetic resonance (EPR) and electron spin resonance (ESR) used on beer stability and antioxidant properties studies (Andersen *et al.*, 2000; Brezova *et al.*, 2002; Kocherginsky *et al.*, 2005). Below, the use of NMR in beer analysis is described in more detail.

1.5.2 Nuclear magnetic resonance (NMR) spectroscopy of beer

High resolution NMR spectroscopy has proven to be a valuable tool in the direct analysis of complex food matrices, such as beer (Duarte *et al*., 2002; Nord *et al*., 2004; Lachenmeier *et al*., 2005; Almeida *et al*., 2006), fruit juices (Cuny *et al*., 2008; Spraul *et al*., 2009) and wine (Viggiani and Morelli, 2008; Son *et al*., 2009; Anastasiadi *et al*., 2009), providing extensive quantitative and structural information on the compositional profile of foods. In fact, NMR spectroscopy, particularly when combined with multivariate analysis (the so called metabonomics approach, described in section 2.5), allows the detection of a large number of compounds simultaneously (in concentrations down to units of mgL^{-1}) offering high throughput and reproducibility conditions, while requiring minimal or no sample preparation, thus allowing potential industrial applications to be envisaged.

Nuclear Magnetic Resonance (NMR) spectroscopy was firstly applied in the study of beer in the 1980s, having involved the use of site-specific natural ratio isotope fractionation (SNIF)-NMR, based on the measurements of isotopic ratios of deuterium/hydrogen (D/H) at specific sites of ethanol and water molecules, for the study of beer quality parameters (Martin *et al.*, 1985; Franconi *et al.*, 1989). The following studies have comprised two main strategies: i) the analysis of beer extracts or raw materials and ii) the direct analysis of beer.

The NMR studies of beer extracts or raw materials were mainly focused on the identification and characterization of specific compounds or family of compounds (targeted approach). In this way, hops acids, iso-α-acids and their derivatives (Pusecker *et al.*, 1999; Nord *et al.*, 2003; Khatib *et al.*, 2007; Intelmann *et al.*, 2009; Intelmann and Hofmann, 2010), carbohydrates (Vinogradov and Bock, 1998), non-starch polysaccharides (Bock *et al.*, 1991; Broberg *et al.*, 2000; Ferre *et al.*, 2000; Dervilly *et al.*, 2002) and beer (Khatib *et al.*, 2006) were characterized. In these studies both 1D

and 2D NMR spectroscopy were applied and, in some cases, combined with other analytical techniques, mainly LC and mass spectrometry (MS).

The direct beer analysis studies by NMR comprised two distinct objectives: the detailed chemical characterization of beer composition and the attempt to assess beer quality parameters. The chemical characterization of beer was performed using both 1D and 2D spectra, allowing the identification of about 30 compounds (Duarte *et al.*, 2002; Almeida *et al.*, 2006). The use of hyphenated methods, LC-NMR and LC-NMR/MS, has also contributed for a more detailed characterization of aromatic composition (Gil *et al.*, 2003) and of carbohydrates (Duarte *et al.*, 2003) in beer. On the other hand, the studies addressing beer quality and the rapid quantification of specific compounds have made use of multivariate analysis (MVA) to interpret large 1D NMR datasets in order to extract meaningful information (metabonomics approach). This NMR/MVA methodology have been applied in an untargeted manner in the differentiation of beer types (top- and bottom-fermentation) and brands (Duarte *et al.*, 2002; Duarte *et al.*, 2004; Nord *et al.*, 2004; Lachenmeier *et al.*, 2005), the assessment and monitoring of beer variability between brewing sites and production dates (Almeida *et al.*, 2006), whereas a targeted approach was followed in the quantification of specific compounds, namely amino and organic acids (Nord *et al.*, 2004) as well as of important brewing parameters, such as original extract and ethanol and lactic acid content (Lachenmeier *et al.*, 2005).

A major issue in NMR metabonomics applied to beer is related with the influence of pH in the obtained NMR spectra. Indeed, NMR spectra are affected by small peak shifts occurring in the metabolites due to small pH variations. The chemical shift parameter is influenced if metabolite exchange between the different chemical states (the protonated/deprotonated states exchange is pH dependent), with the metabolites having slight different resonances frequencies between the different states. If the regime is of fast exchange, the signal will correspond to a single line with a position given by the population weighted average of the chemicals shifts of all the chemical states contributing (Ross *et al.*, 2007; De Graaf, 2007). In this thesis, adjusting sample pH to a relatively low pH value have been found more suitable, taking into account the particular acid mixture present in the sample and the corresponding acidity equilibrium constant, thus minimizing peak shifts, mainly observed for the amino and organic acids resonances (e.g. histidine, acetic and citric acids). Furthermore, the pH adjustment or

buffering to acidic conditions has also been used before in in beer and juices (Nord *et al.*, 2004; Spraul *et al.*, 2009).

In this thesis, the direct analysis of beer by NMR metabonomics was performed to i) identify compositional chemical deviations between beers produced in the same country in order to attain a finer control of production parameters; ii) quantify organic acids, which are important indicators of fermentation performance and of the sensory properties of beer; and iii) characterize the compositional changes occurring during the beer aging process, using a combination of NMR, GC-MS and sensory data, enabling a better understanding of the process and identification of potential new markers.

2. FUNDAMENTALS OF THE ANALYTICAL METHODS EMPLOYED AND OF METABONOMICS

2.1 NMR spectroscopy

2.1.1 Principles of NMR spectroscopy

NMR spectroscopy involves the study of the magnetic properties of nuclei, being based on the effects of an applied external magnetic field on the behavior of nuclei present in the studied molecule. For each observed nuclei, the information provided by NMR depends not only on the kind of nuclei, but also on the electronic environment in which the nuclei are immersed and the positions of nuclei within molecules (chemical environment).

The nuclei observable by NMR are those with a spin quantum number, I, described by positive half integer (1/2) or integer values. When such a nuclei are exposed to the influence of a external field of strength $\mathbf{B_0}$, applied in a direction defined as z, an interaction (Zeeman interaction) occurs between the magnetic moment of a nucleus, $\boldsymbol{\mu}$, and the applied magnetic field $\mathbf{B_0}$. In classical terms, the energy of the nucleus in such a situation is described by the vector product:

$$E = -\boldsymbol{\mu}.\mathbf{B_0} = \mu_z.\mathbf{B_0} = -(\gamma.\hbar.m_I).\mathbf{B_0} \qquad \text{[Eq. 2.1]}$$

where μ_z is the component of $\boldsymbol{\mu}$ along the direction of the applied field $\mathbf{B_0}$, γ refers to the magnetogyric ratio of the nucleus, m_I is the magnetic quantum number representing the different orientations of $\boldsymbol{\mu}$ relatively to the applied field $\mathbf{B_0}$, and the quantity $\hbar$ is the Planck's constant, h, divided by 2π. In the presence of $\mathbf{B_0}$, the normally degenerate energy of the nucleus becomes non-degenerate giving (2I+1) energy levels for a nucleus with spin quantum number I, corresponding to 2I+1 values of m_I. For example, in magnetic nuclei with spin I = 1/2 (e.g. ^{1}H, ^{13}C, ^{15}N, ^{19}F), there are two possible spin states: +1/2 (α-state) and -1/2 (β-state), the former being of lower energy, as seen in Figure 2.1.

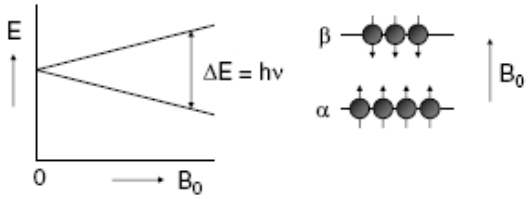

Figure 2.1. Energy levels for a nucleus of spin quantum number I = 1/2, as a function of external magnetic field strength $\mathbf{B_0}$ (reproduced from De Graaf, 2007).

As there are only two energy levels possible (when I= 1/2), therefore, the energy difference, ΔE, between the quantized states α- and β-state is given by:

$$\Delta E = |\gamma.\hbar.\mathbf{B_0}| = h.\upsilon \qquad \text{[Eq. 2.2]}$$

In terms of classical mechanics, the field imposes a torque on the moment **μ**, which results in a circular motion along the axis of the applied field $\mathbf{B_0}$. This motion is referred to as Larmor precession and occurs at a rate $\omega_0 = -\gamma.B_0$ ($rads^{-1}$) or $\upsilon = -\gamma.B_0/2\pi$ (Hz). For the resonance condition to be satisfied, the frequency of the electromagnetic radiation applied must match that of the Larmor precession, i.e. the resonant frequency of a nucleus is simply its Larmor frequency (Claridge, 1999; Lambert and Mazzola, 2003). At thermal equilibrium, there will be an excess of nuclei in the α-state as defined by Boltzmann distribution:

$$P_\beta/P_\alpha = \exp(\Delta E/k.T) = \exp(h.\gamma.\mathbf{B_0}/2\pi.k.T) \qquad \text{[Eq. 2.3]}$$

where P_β and P_α represent the spin populations in the β and α energy levels, respectively, k is the Boltzmann constant and T is system temperature. Because ΔE values are relatively small, only about 1 part of 10^4 are at the highest available field strengths, justifying the intrinsic low sensitivity of NMR when compared with other analytical techniques, which is further enhanced by the low natural abundance of certain nuclei, such as ^{13}C (approximately 1.1%). The small population excess of nuclear spins can be represented as a collection of spins distributed randomly about the precessional cone and parallel to the *z*-axis, showing no net magnetization in the transverse (*x*-*y*) plane, Figure 2.2, giving rise to a resultant bulk magnetization vector, $\mathbf{M_0}$, along the z-axis. This simplified picture is generally referred to as the vector model of NMR or as the Bloch vector model.

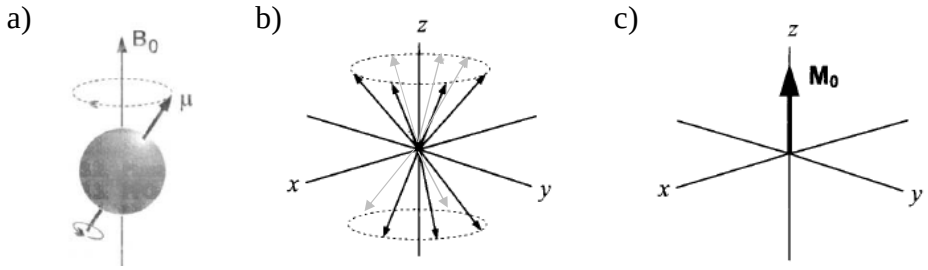

Figure 2.2. a) Precession of a magnetic nucleus caused by the applied static magnetic field $\mathbf{B_0}$; b) two precessional cones for a collection of 1/2 spin nuclei in the α- and β-states; and c) net spin magnetization $\mathbf{M_0}$ at equilibrium (reproduced from Claridge, 1999).

When the sample is subject to a short pulse (few μs) of radiofrequency (rf) irradiation, with a magnetic component $\mathbf{B_1}$, to excite all frequencies of a given nucleus

at the same time, the net magnetization $\mathbf{M_0}$ is shifted toward the transverse (*x*-*y*) plane, where it can be electronically detected, as illustrated in Figure 2.3. To help visualize the effects of the oscillating field $\mathbf{B_1}$ on the bulk magnetization vector, a simplified representation is employed, known as the rotating frame of reference. This representation consists in using a reference frame (*x'*, *y'*, *z'*) rotating at a frequency ω_0 around the *z*-axis, so that the $\mathbf{B_1}$ vector appears static (Claridge, 1999). The angle θ through which the magnetization vector **M** turns is dependent on the amplitude and duration of the pulse. For example, if the rf is turned off just as the **M** vector reaches *y*-axis, this represents a 90 º pulse and maximum signal intensity is observed. On the other hand, if **M** vector becomes oriented along the *z*-axis, representing a 180 º pulse, no signal is detected, since only magnetization on the (*x*-*y*) plane is able to induce a signal in the detection coil (Claridge, 1999).

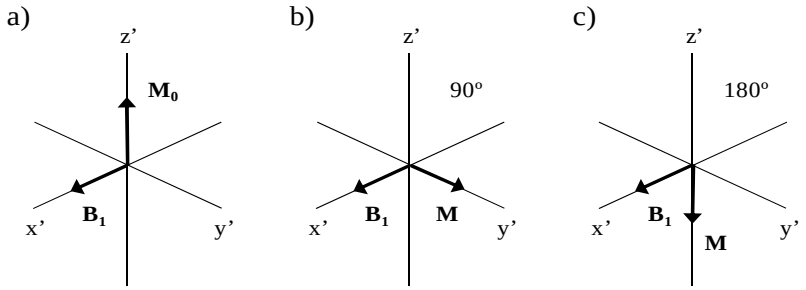

Figure 2.3. Representation of a) the net magnetization $\mathbf{M_0}$ under equilibrium conditions; and the effects on the $\mathbf{M_0}$ of a b) 90º and c) 180º rf pulses.

Immediately after the pulse, the system will adjust to re-establish the Boltzmann distribution, and so the transverse **M** (magnetization in the *x*-*y* plane) will gradually disappear and, simultaneously, grow along the z-axis towards it initial value and direction ($\mathbf{M_0}$). This return to equilibrium, referred to as relaxation, causes the NMR signal to decay with time, producing the observed Free Induction Decay (FID). A mathematical calculation, the Fourier Transform (FT), is then applied to transform the time domain data into the frequency domain (common NMR spectrum). Relaxation processes can be divided into two categories: i) longitudinal or spin-lattice relaxation (recovery of magnetization along the z-axis) corresponding to the re-establishment of the equilibrium of populations and involving a energy transfer between the spins and the surroundings (lattice); and ii) transverse or spin-spin relaxation (decay of magnetization in the x-y plane) involving an energy transfer between the magnetized spins. Both processes follow an exponential behavior, occurring simultaneously but at different

rates, denoted by T^{-1} and T^{-2}, respectively. The relaxation processes are induced by field fluctuation due to molecular motion, with the relaxation times characteristics of each process (T_1: spin-lattice relaxation and T_2; spin-spin relaxation) closely related with the molecular mobility, characterized by a rotational correlation time τ_c (average time taken for the molecule to rotate through 1 radian) (Claridge, Lambert and Mazzola, 2003).

There are several useful parameters that can be extracted from NMR spectra, providing important information for molecular structure characterization: chemical shift (δ, expressed in part per million, ppm), spin-spin coupling pattern (multiplicity), coupling constant (*J*) and signal intensity.

Nuclei are surrounded by electrons, which shield them from $\mathbf{B_0}$. Thus, the effective magnetic field experienced by the nucleus, $\mathbf{B_{eff}}$, results from the contributions of $\mathbf{B_0}$ and the local magnetic fields produced by the surrounding electrons (since $\mathbf{B_o}$ induces currents in the electrons clouds), as expressed in Eq. 2.4:

$$\mathbf{B_{eff}} = \mathbf{B_0}\,(1-\sigma) \qquad \text{[Eq. 2.4]}$$

where σ is the shielding constant, which reflects the extent to which the electron cloud around the nucleus shields it from the external magnetic field. In the case of protons, they are magnetically shielded to different extents depending on their chemical environment and location (i.e. their neighbouring atoms and type of chemical bond). Those slightly different magnetic fields experienced by each site will correspond to a different position on the NMR spectrum. The position of a resonance signal in a NMR spectrum is measured by its resonance frequency (υ_{sample}, expressed in hertz (Hz)), being this value dependent of the strength of the applied $\mathbf{B_0}$. To standardize the frequency scale of NMR, the obtained resonance frequencies are i) normalized using a reference compound (set to 0.0 ppm); and ii) divided by the spectrometer frequency (in MHz) and multiplied by a factor of 10^6. The resulting dimensionless value is named chemical shift, δ, and is given in parts per million (ppm), which is, therefore, constant regardless of the operating frequency of the spectrometer. Examples of some typical range of chemical shifts values for ^{1}H and ^{13}C are listed in Table 2.1.

The interaction between neighbouring magnetic nuclei within a molecule is referred to as spin-spin coupling or scalar coupling, which causes the splitting of NMR signals into multiplets. The multiplicity of a signal (i.e. the number of peaks in the split signal) is giving by (2nI+1), where n refers to the number of neighbouring equivalent nuclei.

Table 2.1. Characteristic chemical shifts (δ) range for ^{1}H and ^{13}C in common chemical groups (reproduced from Fan, 1996).

^{1}H chemical shifts		**^{13}C chemical shifts**	
Chemical group	**δ/ ppm**	**Chemical group**	**δ/ ppm**
C−CH_3	0.7-1.1	−CH_3 (methyl)	8-30
=C−CH_3	1.5-1.8	−CH_2 (methylene)	14-55
$COCH_3$ (acetyl)	2.0-2.5	−CH− (methine)	22-60
N−CH_3	2.5-3.3	−C− (quaternary)	30-40
O−CH_3	3.3-4.3	CH_2=C−C−R	100-150
C−CH_2−C	1.3-2.5	C=C−C=C−R	110-150
O−CH_2−C	4.0-5.0	Heteroaromatic ring	100-165
C−CH−OH	3.3-4.0	C−OH (alcohol)	44-85
C−CH−O−ester	4.2-5.3	C−O−C (ester)	55-85
=CH−C (olefinic)	5.5-8.5	R−COOH (saturated)	165-188
Aromatic ring	6.0-9.0	R−C=C−COOH	158-174
−CHO (aldehyde)	9.0-10.2	R−COOR' (saturated)	158-178
−COOH	10.5-13.5	R−C=C−COOR'	152-172
C−OH (alcohol)	1.5-6.0	R−CHO (saturated)	196-220
C−OH (phenol)	6.5-18.5	R−C=C−CHO	176-195
Primary amines	1.1-1.8	Saturated ketones	195-220
Secondary amines	1.2-2.1	C−N−R_2	20-70
CO−NH	5.0-6.5	R−CO−NH_2	150-178

For ^{1}H NMR spectra of isotropic liquids, under appropriate conditions, the area of each peak is directly proportional to the number of the corresponding nuclei (Rizzo and Pinciroli, 2005). In this context, for quantitative NMR acquisition conditions, the use of inter-scan times of at least 5 x highest T_1 is necessary for all compounds under study, allowing the complete spin-spin relaxation of all spins (Evilia, 2001). Although ^{1}H NMR is the most commonly used for quantification of analytes in complex matrices, ^{13}C NMR has also been used, providing complementary structural information while reducing the problem of overlap usually occurring in ^{1}H NMR. However, ^{13}C NMR suffers from poor sensitivity due to low natural abundance (~1.1%, when compared with 99.97% of ^{1}H) and low gyromagnetic ratio of ^{13}C nuclei (~25% of ^{1}H) (Shaykhutdinov *et al*., 2009). For quantification by signals integration, the selection of an appropriate standard reference compound is required, the most commonly used in complex matrices being 3-(trimethylsilyl)propionate (TSP), although interactions between TSP and proteins have been detected (Shimizu *et al.*, 1994). An alternative is the use of the Electronic Reference To access In vivo Concentrations (ERETIC) method, based on an electronic reference signal, and thus improving quantification precision (Barantin *et al.*, 1997; Akoka and Trierweiller, 2002; Rizzo and Pinciroli, 2005).

Furthermore, recent improvements allow the routine use of fully automated NMR systems (Spraul *et al.*, 2009), thus reducing the variability/uncertainty associated to the need of human intervention in quantitative NMR, particularly in spectra processing and in the choice of the integrated signal and integral tail settings.

2.1.2 Two-dimensional (2D) NMR spectroscopy

The first approach generally used to assign the NMR signals of complex mixtures consists in the measurement of chemical shifts, multiplicity and coupling constants in 1D spectrum. However, high spectral complexity encountered in those mixtures poses the problem of strong signal overlap in all spectral regions, thus hindering signal assignment. 2D NMR spectroscopy provide a valuable approach to carry out spectral assignment, as the information is spread in two dimensions and spectral overlap is significantly reduced. Independently on the nature of the interaction to be depicted, all 2D NMR experiments have the same basic format (Figure 2.4), being sub-divided into four steps (Jacobsen, 2007):

i) Preparation: excite nucleus **A**, creating magnetization in the x-y plane.
ii) Evolution: indirectly measure the chemical shift of nucleus **A**.
iii) Mixing: transfer magnetization from nucleus **A** to **B**.
iv) Detection: measure the chemical shift of nucleus **B**.

The preparation and mixing steps typically comprise a pulse or a cluster of pulses and/or fixed periods of time, depending on the nature of the experiment. The key step for the generation of the second dimension is the evolution step. This is accomplished by simply waiting for a period of time (called t_1, the evolution period) after which the spin system is altered. The resonance signal is then recorded during the subsequent detection time, t_2. The experiment is repeated, with the delay t_1 incremented each time, whereby the initial state of the nuclear spin system during the detection period at $t_2 = 0$ changes accordingly. After data recording, all FID's are transformed with respect to t_2 to obtain a set of spectra in which the peaks intensities or phases are modulated as a function of the delay t_1. Fourier Transform (FT) with respect to t_1 converts the frequency modulation into peaks in the two-dimensional spectrum, usually represented as a contour plot.

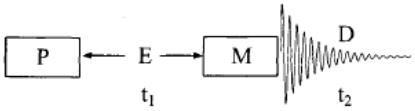

Figure 2.4. General scheme for 2D NMR experiment (adapted from Claridge, 1999).

There are, essentially, two categories of 2D experiments: i) correlated experiments, in which the resonance frequency of one peak is related to those of its neighbours and thus molecular connectivities or distances between atoms can be determined; and ii) resolved experiments (based on *J*-coupling interaction), in which the frequency axes show two different interactions (e.g. coupling constant information and chemical shift data). Furthermore, the 2D experiments can be sub-divided into homonuclear (nuclei from the type, usually ^{1}H) and heteronuclear (two different types of nucleus, e.g. ^{1}H and ^{13}C) experiments. Below the used 2D NMR experiments in this thesis are described, namely total correlation spectroscopy (TOCSY), heteronuclear single quantum correlation (HSQC) and *J-resolved* experiments.

The TOCSY experiment (Braunschweiler *et al.*, 1983) correlates homonuclear scalar-coupled spins, usually ^{1}H, that sit within the spin-system, regardless of whether they are themselves coupled to one another. The TOCSY sequence, in this case, the MLEV-17 spin-lock sequence (Malcolm Levitt´s composite-pulse decoupling sequence) (Figure 2.5), allows the magnetization to propagate along continuous chains of spin-spin coupled protons by relaying coherence from one proton to the next along the chain. The requirement for nuclei to experience identical local fields for magnetization transfer to occur between them is referred as the Hartmann-Hahn match (Claridge, 1999; Jacobsen, 2007). Long mixing periods will enable magnetization to travel further along the nuclei chain and peaks will arise from multi-steps transfers. However, the use of very long mixing periods could lead to sensitivity loss due to relaxation effects (in the rotating frame) of the spin-locked magnetization. Since magnetization may travel in either direction along the spin-chain, 2D TOCSY spectra are symmetrical about the diagonal.

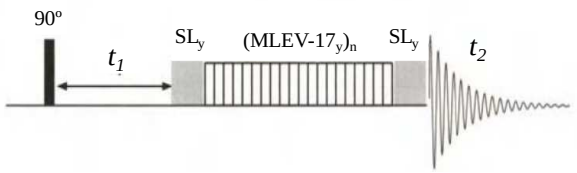

Figure 2.5. TOCSY pulse sequence using the MLEV-17 spin-lock sequence (Malcolm Levitt´s composite-pulse decoupling sequence) (reproduced from Claridge, 1999).

The HSQC experiment (Figure 2.6) enables correlating heteronuclear spins across a single bond, hence identifying directly connected nuclei (commonly ^{1}H-^{13}C) and employ detection of the high-sensitive nucleus (e.g. ^{1}H, ^{19}F, ^{31}P instead of ^{13}C). In the HSQC experiment (Bodenhausen and Ruben, 1980) only single-quantum transverse magnetization of the heteronuclear spins evolves during the t_1 (Figure 2.6), being the

transverse heteronuclear magnetization generated by polarization transfer from the attached proton via the insensitive nuclei enhanced by the polarization (INEPT) sequence. Consequently, the crosspeaks in the HSQC spectrum do not contain homonuclear ^{1}H-^{1}H coupling along the f_1. This results in an improved resolution in this dimension, especially for small molecules (Keniry and Sanctuary, 1992). However, this also leads to a greater number of pulses utilized, promoting losses from rf inhomogeneity, pulse miscalibration or off-resonance excitation.

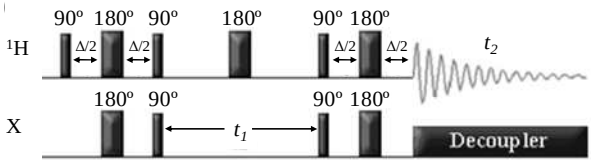

Figure 2.6. HSQC pulse sequence (reproduced from: http://www.chem.queensu.ca).

The *J*-resolved method aims to separate chemical shifts from scalar coupling parameters (chemical shifts in f_2 and spins coupling in f_1 dimensions), and they can be homonuclear or heteronuclear experiments. In this thesis, homonuclear experiments have been employed. As seen in Figure 2.7, the homonuclear *J*-resolved experiment (Aue *et al.*, 1976) employs a spin-echo sequence during the evolution period t_1, in order to make the detected FIDs insensitive to chemical shift evolution during t_1 period, so that the f_1 dimension only contains coupling information. On the other hand, the f_2 dimension contains both chemical shift and coupling constant information, because one cannot simultaneously broadband decouple and observe the ^{1}H spectrum. Thus, rather than lying parallel to f_1 axis, the ^{1}H multiplets sit at 45º to either axis (assuming identical plot scaling for both dimensions). Therefore, the columns parallel to f_1 do not display the expected ^{1}H multiplets and the f_2 projection displays both chemical and scalar couplings. To overcome these shortcomings and retaining chemical shifts in f_2, post-processing techniques are commonly applied, namely by tilting the multiplets through an angle of 45º about their midpoints (Kadkhodaie *et al.*, 1991).

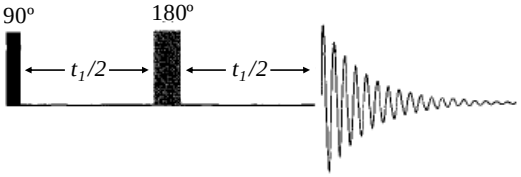

Figure 2.7. Homonuclear *J*-resolved pulse sequence (reproduced from Claridge, 1999).

2.1.3 Hyphenated NMR Methods: LC-NMR(/MS)

Liquid chromatography (LC) is one of the most applied techniques for the analysis of complex mixtures. Conventional detectors used (e.g. refractive index, ultra violet (UV), fluorescence and electrochemical detectors) show limited ability to obtain components structural information. This led to a growing interest in the hyphenation of LC with techniques that provide detailed structural information, such as NMR spectroscopy and mass spectrometry (MS). In fact, the hyphenated LC-NMR and LC-NMR/MS methods have been applied in complex mixtures studies, such as in beer (Gil *et al*., 2003; Duarte *et al*., 2004), fruit juices (Duarte *et al*., 2005) and wine (Gil *et al*., 2003), as they provide complementary structural information, increasing the capability of unambiguous identification of detected components.

Figure 2.8 shows a schematic representation of a typical set-up for LC-NMR(/MS), consisting in an injection device, LC pumps together with a gradient unit, a LC column and a UV detector (LC part), a NMR spectrometer, equipped with a dedicated flow probe, and a mass spectrometer, attached to the system via a splitter at the output of the LC-NMR interface (Albert, 2002). NMR and MS are coupled in parallel so it is possible to choose the time for MS detection (i.e. prior, after or simultaneous with NMR detection).

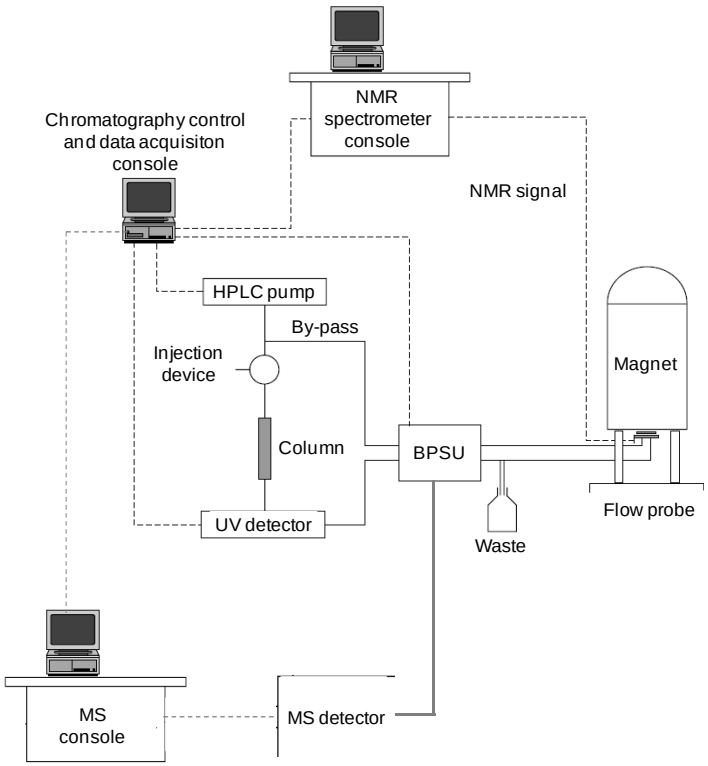

Figure 2.8. Schematic representation of a LC-NMR(/MS) system; (----) refers to electronic junctions; (___) refers to capillary junctions (adapted from Albert, 2002).

The heart of the LC-NMR(/MS) is the flow control, which determines the working mode of the apparatus. There are four main operation modes, with both isocratic and gradient elution: i) on-flow, ii) direct stop-flow, iii) time-slice, and iv) loop storage (Albert, 2002; Elipe, 2003). In the on-flow mode, the NMR and MS data are acquired continuously while the sample is flowing through the detection cell. This method allows the rapid screening of the analytes, however with NMR sensitivity limited by the short acquisition time. Moreover, the use of flowing gradient elution originates inhomogeneities of the magnetic field, leading to deterioration in the spectral resolution, and shifts of the solvent and samples resonances.

In the direct stop-flow mode, a peak is selected from the chromatographic run, which is stopped when the LC peak has reached the NMR flow cell, allowing 1D and 2D NMR measurements to be performed with longer acquisition times. After the NMR measurements, the LC separation is continued until the next peak is positioned in the NMR flow cell. The flow may be stopped manually or automatically using the time-slice method, which consists of stopping the flow at short intervals over the chromatogram to time-slice different parts. This method may be useful if poor chromatographic separation occurs, if the relevant compounds under investigation have weak UV chromophores or if the exact chromatographic retention time is poorly defined. However, when operating in stop-flow mode, the chromatographic resolution is degraded due to diffusion that will broaden the peaks remaining in the chromatographic system. Furthermore, no MS data can be acquired because of the interruption of the flow during the NMR measurements.

In loop storage mode, the selected peaks directly flow from the chromatographic system into the storage device (i.e. storage loops) without the need of interrupting the chromatography, allowing to place the remaining fractions of interest in storage loops for subsequent analysis. After the chromatographic run is finished, the loop contents may be transferred sequentially to the NMR and MS spectrometers for off-line experiments. Once the peaks are stored into the loops, the NMR acquisition time are not limited by no diffusion effects, therefore, this operation mode overcomes the disadvantages of the stop-flow methods (Albert, 2002).

Advances in LC-NMR technique, such as the development and optimization of dedicated flow probes, have improved greatly the physical connection between LC and NMR. Furthermore, the combined hyphenation of both NMR and MS to the separation technique of LC, enabled obtaining complementary information for each analyzed

fraction and becoming extremely useful tool for unambiguous structural elucidation (Elipe, 2003; Exarchou *et al.*, 2005).

2.2 Capillary Electrophoresis (CE)

CE is a separation technique based on the differential movement or migration of ions based on charge, when subjected to an electric field (Altria, 2000; Frazier *et al.*, 2000). In practice, separation of analyte ions, anions (negatively charged) and cations (positively charged), is performed in an electrolyte solution (the background electrolyte, BGE) present in a narrow fused-silica capillary (normally ranging from 25 to 100 µm in internal diameter, ID).

The instrumentation required for CE is illustrated in Figure 2.9. The ends of the capillary are placed in separate buffer (electrolyte) reservoirs, each containing an electrode, a positive (anode) and a negative (cathode), connected to a high-voltage power supply. The sample is injected onto the capillary by applying pressure (hydrodynamic injection) or voltage (electrokinetic injection) for a few seconds After that, an electrical potential (between 5-30 kV) is applied across the capillary and the separation is performed, with the formation of zones of analyte, depending on the different electrophoretic mobilities of ionic species, and migrate towards the outlet side of the capillary, into the electrode of opposite charge. Optical (e.g. UV-visible or fluorometric) detection can be achieved directly through the capillary wall near the opposite end.

Separation by CE relies on the differences in the speed of migration (migration velocity) of analytes, expressed as:

$$\upsilon = \mu_e.E \qquad \text{[Eq. 2.5]}$$

where υ is ion migration velocity (ms^{-1}), μ_e refers to the electrophoretic mobility ($m^2 V^{-1} s^{-1}$) and E is the electric field strength, a function of the applied voltage divided by the total capillary length, ($V\ m^{-1}$). Electrophoretic mobility indicates how fast a given analyte may move through a given environment (buffer solution), reflecting the balance of forces acting on each individual analyte: the electrical force acts in favor of motion whereas the frictional force acts against it. Since these forces are in steady state during electrophoresis, the electrophoretic mobility is constant (for a given analyte, under a set of conditions) and it is given by:

$$\mu_e = q / (6.\pi.\eta.r) \qquad \text{[Eq. 2.6]}$$

where q refers to the charge on the analyte, η is the solution viscosity, and r corresponds to the analyte radius. The charge of the analyte is fixed for fully dissociated ions (e.g. strong acids, small ions), but can be affected by pH changes in the case of weak acids and bases. Furthermore, variations in the electrophoretic mobility can be caused by differences in the charge-to-size ratio of analyte ions, with higher charged, smaller size ions conferring greater mobility (Frazier *et al.*, 2000).

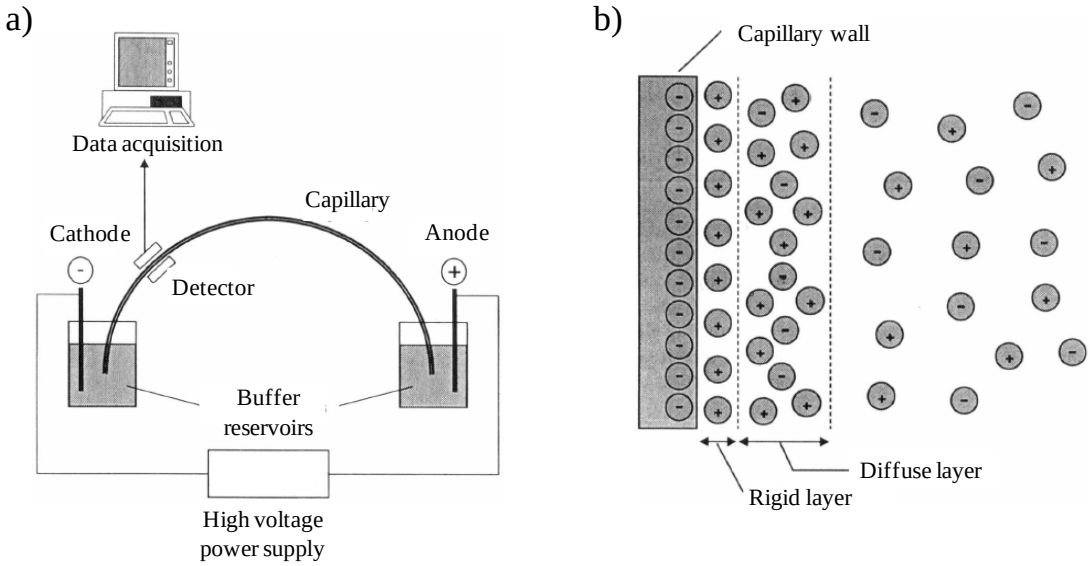

Figure 2.9. a) Schematic representation of a typical CE system; b) Stern's model of the double-layer charge distribution at a negatively charged capillary wall (reproduced from Frazier *et al.*, 2000).

A critical factor in CE is the bulk flow of liquid through the capillary, named electroosmotic flow (EOF). EOF reflects the fact that, when filled with buffer, the negatively charged capillary tube walls (due to silanol groups dissociation) attract positively charged ions from the buffer solution, originating an electrical double layer and a potential difference (zeta potential, ζ) close to the capillary wall as described accordingly to Stern's model (Figure 2.9b). Applying a voltage across the capillary, cations in the diffuse layer are free to migrate, carrying with them the bulk solution, originating a net flow in the direction of the cathode, with a velocity, υ_{EOF}, determined by:

$$\upsilon_{EOF} = [(\varepsilon_0.\varepsilon.\zeta)/(4.\pi.\eta)].E \qquad \text{[Eq. 2.7]}$$

where ε_0 and ε are, respectively, the dielectric constant of a vacuum and the buffer, and η is the buffer viscosity (the mobility of the EOF,μ_{EOF}, is determined by the terms enclosed in square brackets). As EOF has a flat profile (i.e., its charge on the capillary

wall is uniformly distributed along the capillary), no pressure drops are encountered and the flow velocity is uniform. This contrasts with the pressure-driven flow (as in the case of LC technique) that, by the use of an external pump, yields a laminar or parabolic flow profile, as seen in Figure 2.10. This is a huge advantage of CE technique, as it minimizes zone broadening, leading to high separation efficiency that allow separations on the basis of mobility differences as small as 0.05%.

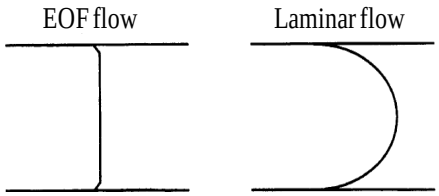

Figure 2.10. Cross sectional flow profile of EOF and laminar flow (adapted from Frazier *et al.*, 2000).

The data output of CE is analogous to a chromatogram and it is named electropherogram. An electropherogram is a plot representation of the detector response as a function of migration time, being the UV-visible absorbance or fluorescence the most used detectors. In UV detection (detector used in this thesis), two types of detection modes can be used: direct and indirect detection. In direct detection, a single signal wavelength is chosen for the monitoring of the analytes migration. In indirect detection, the analytes displace a UV-absorbing component, the BGE, resulting in a negative signal which is reversed by monitoring a signal wavelength outside the absorption area of BGE (the latter cancelled by a reference wavelength).

2.3 Gas Chromatography–Mass Spectrometry (GC-MS)

GC is based on the physical separation method in which the components in a complex mixture are selectively distributed between the mobile phase (an inert carrier gas) and a stationary phase (coating of column packaging or of the inner column wall). The chromatographic process consists in repeated sorption/desorption steps, occurring during the movement of the analytes along the stationary phase caused by the carrier gas (Karasek and Clement, 1986). The resulting separation corresponds to differences in the distribution coefficients of the individual analytes, K_{fs}, which can be defined as the ratio between analyte concentration present in stationary, C_s, and mobile, C_m, phases:

$$K_{fs} = (C_s / C_m) \qquad \text{[Eq. 2.8]}$$

The retention time corresponding to an analyte i, $t_{r,i}$, can be expressed as the time distance between the sample injection and the top of the peak due to the analyte. From this, a number of critical GC parameters can be calculated (Table 2.2), namely: mass distribution, k'_i, separation factor, $\alpha_{j,i}$, number of theoretical plates, N, peak width and peak width at half weight, W and $W_{1/2}$, respectively, and resolution, $R_{i,j}$ (Rood, 2006).

Table 2.2. Some GC parameters (Rood, 2006).

Parameter	**Equation**	
Capacity ratio or mass distribution	$k'_i = (C_s.V_s)/ (C_m.V_m)$	[Eq. 2.9]
Separation factor	$\alpha_{j,i} = (k'_j/ k'_i) \geq 1$	[Eq. 2.10]
Number of theoretical plates	$N = 5.545\ (t_r / W_{1/2})^2$	[Eq. 2.11]
Height equivalent to a theoretical plate	$H = L/N$	[Eq. 2.12]
Resolution	$R_{i,j} = [2(t_{r,j} - t_{r,i})/(W_j + W_i)]$	[Eq. 2.13]

The column efficiency is gauged by the number of theoretical plates, *N*, i.e. higher *N* corresponds to higher potential of the column to separate two peaks. Since *N* depends on the length of the column, column efficiency can also be expressed by *H* (Eq. 2.11), where L refers to the length of the column. The effect of flow in the column is usually shown by plotting *H* as a function of flow rate or linear velocity, u, as indicated by Van Deemter equation (Eq. 2.14):

$$H = A + (B/u) + C.u \qquad \text{[Eq. 2.14]}$$

where A, B and C are constants, with A representing the contribution of eddy diffusion, B of longitudinal diffusion, and C (composed by C_m and C_s) resistance to mass transfer in mobile and stationary phases, respectively. Resolution is a quantitative parameter related to the separation between two adjacent peaks, which is determined considering the retention times and the peak width of the adjacent peaks *i* and *j* (Eq. 2.13) and, thus depends on the plate height, column length, separation factor and capacity ratio.

The GC instrumentation consists, essentially, of a sample introduction device (injector), a column housed in the temperature-programmable oven, an interface, the mass spectrometer and a data collecting system (McMaster, 2008). The most widely applied GC injector is the split/splitless injector (Figure 2.11a). There are three main types of injection: split, splitless and on-column. In split injection, only a small fraction

of the sample (typically ranging from 1:10 to 1:100, depending on the column internal diameter) will be sent to the column, avoiding overloading the low-volume capillary columns. This is used when in analytes present in high concentrations. In splitless injection, the splitter vent is closed for a period of time, while the sample and solvent flows onto the top of the column. After that, the splitter vent is opened, concentrating them at the top of the column. Splitless injection has the benefit of higher sensitivity, although can be troublesome when used in combination with capillary columns, due to column's low sample volume allowed. This type of injection is particularly applied in samples with trace amounts. Finally, in on-column injection the sample is directly introduced into the column, being useful for analyzing components with high boiling point and/or for analytes that could degrade at high temperatures.

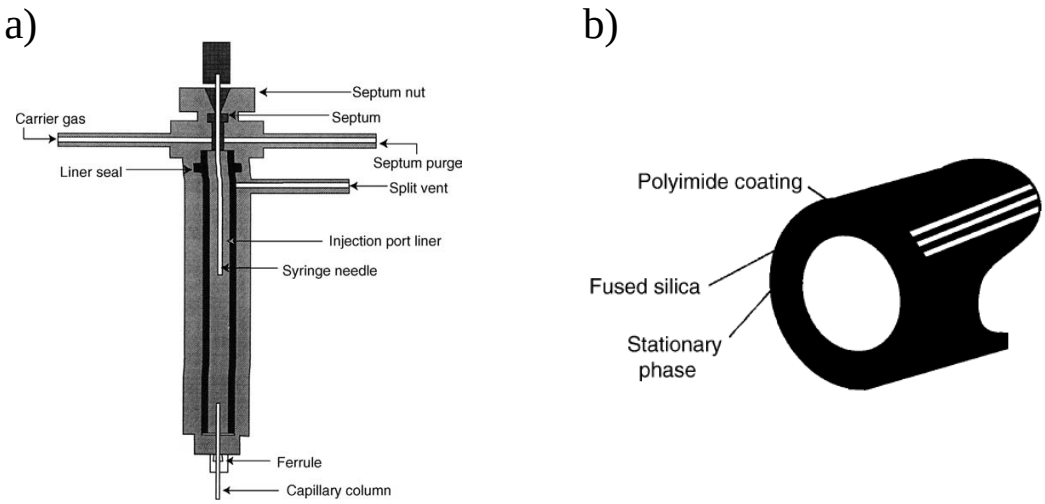

Figure 2.11. a) Typical split/splitless GC injector; b) representation of a typical capillary column, transversal sight (adapted from McMaster, 2008).

The most commonly applied GC columns are capillary columns. A wide range of capillary columns are available, differing in length, inner diameter, film thickness and type of stationary phase. The typical capillary columns used for GC are constituted, normally, by three distinct parts: the stationary phase, the fused silica tube and the polyimide coating (Figure 2.11b). In the selection of the stationary phase, a similar polarity to the analytes of interest is generally selected. In GC, two separation mechanisms may be exploited: i) separation according to boiling points of the solutes (especially when using apolar stationary phase) and ii) separation based on selective partitioning, i.e. interactions of analytes with the stationary phase (prevalent when using polar columns).

The GC column is housed in a thermostatted oven, allowing the column to operate at a constant temperature (isothermally) or, more commonly, using a temperature program, the latter enabling obtaining narrow peak widths and good separation for both early eluting as well as the more retained compounds (Jennings *et al.*, 1997). To sweep the analytes through the column, a carrier gas at a given pressure is applied, being the most used: hydrogen, helium or, to a lesser extent, nitrogen. The type of carrier gas and its flow velocity have a relevant effect on the chromatographic separation, with the optimum flow value being determined by applying the van Deemter equation, mentioned before. The analytes that are eluted from the column have to be detected by a proper device, a detector. Several detectors have been used in tandem with GC, being the most common the flame ionization detector (FID), the electron capture detector (ECD) and the mass spectrometer (MS). In fact, due to its versatility, robustness and sensitivity, MS is nowadays one of the most valuable detectors available. As MS detector was the only detector used in the work developed in this thesis, it will be discussed in detail.

The principle of MS is the production of gas-phase ions that are subsequently separated according to their mass-to-charge (*m/z*) ratio. The resulting mass spectrum is a plot of the (relative) abundance of the generated ions as a function of the *m/z* ratio. A mass spectrometer essentially consists of five parts: sample inlet system, an ion source (where ionization and fragmentation takes place), a mass analyzer (for separation of the ions according to their *m/z* ratio), an ion detector, commonly electron multiplier, and a system of data handling.

The sample is introduced into the MS directly or via an open split coupling. The ion source consists of a filament providing high energy electrons for ionization, and various lens for guiding the ions into the analyzer. A wide variety of ionization techniques are available, being the most common the electron ionization (EI) and the chemical ionization (CI). In EI, the analyte vapor is subjected to bombardment by energetic electrons, typically of 70 eV, which is substantially above the ionization potential of most organic molecules and, therefore, sufficient to cause both ionization and fragmentation. In CI, a reaction gas (such as methane, isobutene or ammonia) is ionized by energetic electrons from the filament and the resulting ions ionize sample analytes by charge transfer process. The result is a “softer” ionization technique with

lesser fragmentation extent and a higher possibility of obtaining molecular ion indicative of the compound's molecular weight.

Mass analysis (i.e. separation of ions according to their *m/z* ratio in either time or space) can be achieved by various types of mass analyzers, being the most common the quadrupole, ion trap, the time of flight and magnetic sector instruments. Following, the basic principles of the ion trap are described, as it was the analyzer used in thesis. An ion trap consists of a cylindrical ring electrode, to which a quadrupole field is applied, and two end cap electrodes (Figure 2.12). The mass analysis process is a discontinuous, pulsed process consisting of a number of consecutive steps. Firstly, in the pre-ionization step, a short ionization pulse is generated and the ions are removed from the ion trap. After setting an appropriate rf storage voltage at the ring electrode, ions are generated by EI either inside the trap or injected from the external ion source and stored. In full-scan mode, the ions of different *m/z* are consecutively ejected by ramping the rf voltage at the ring electrode, with the ejected ions being detected by an electron multiplier outside the trap. The fragments that pass the analyzer strike the surface of the detector after first being deflected away from a straight path out of the analyzer by a lens (amu offset). The most used detector is the electron multiplier, which consists of various dynodes, each of which release numerous electrons when hit an ion or electron. In this manner, a cascade effect is produced to deliver a gain in the signal in the order of 10^6, depending in the applied voltage. A collector plate, placed on the last dynode, is connected to a preamplifier that converts the current into a voltage, suitable to be measured. That resulting signal, voltage, is sent to the system of data handling where the information is processed.

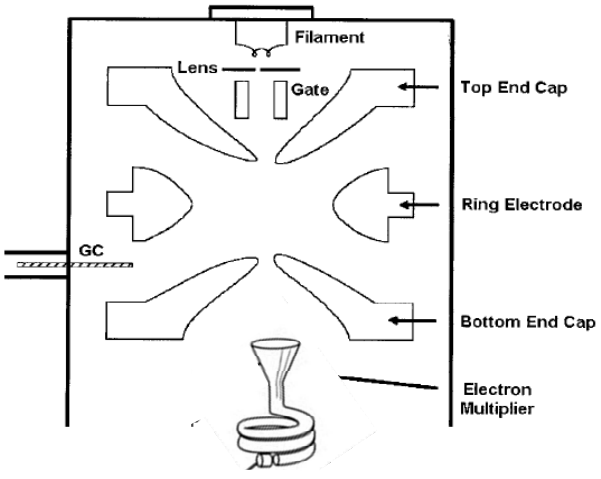

Figure 2.12. Schematic diagram of an ion trap detector (adapted from Niessen, 2001).

Quantitative analysis by GC-MS is based upon the fact that the area or height of a chromatographic peak is proportional to the quantity of the compound injected, within the dynamic range of the detector. In order to determine the accurate concentration of a certain analyte, several techniques can be used, with special attention to the internal standard (IS) methodology. In IS method, the area or height of a IS peak (compound added to the sample at a specific concentration) is used to normalize those of the analytes of interest, correcting small variations that may occur in peak area/height due to injection or minor operational oscillations. For an accurate, absolute analytes concentration determination, the buildup of calibration curves is needed. The correct selection of the IS compound is crucial, being important for the selected IS compound to be chemically and physically similar to the analytes, chemically inert, have a similar retention time and concentration level than the analytes and being well-resolved from all the other chromatographic peaks arising from the sample (Grob, 1995).

2.4 Fourier Transform-Infrared (FT-IR) Spectroscopy

Infrared (IR) spectroscopy is a technique based on the vibrations of the atoms of a molecule. The IR spectroscopy deals with the infrared region of the electromagnetic spectrum, being that region divided into: i) the higher energy near-IR (NIR), approximately between 14000-4000 cm^{-1}, that can excite harmonic vibrations; ii) the mid-IR (MIR), between 4000-400 cm^{-1}, that may be used to study the fundamental vibrations and associated rotational-vibrational structure; and ii) the far-IR (FIR), approximately 400-10 cm^{1}, lying adjacent to the microwave region and may be used for rotational spectroscopy (Stuart, 2004). Particularly, MIR spectroscopy involves the molecular absorption of radiation arising through transitions between vibrational energy states and rotational sub-states of the molecule, the latter not quantized in liquids or in solids. In the absorption process, those frequencies of IR radiation that match the vibrational frequencies of the molecule are absorbed, with that energy being used to increase the amplitude of the vibrational motions of the bonds in the molecule. However, only those bonds that have a dipole moment that changes as a function of time are capable of absorbing IR radiation (Stuart, 2004).

The positions of the absorption bands seen in the MIR spectrum depend on the masses of the atoms and on the force constant of the vibrating bonds. Thus, many functional groups in organic molecules show characteristic vibrations, which correspond to absorption bands located in small portion regions of the IR spectrum.

Typical absorption frequencies bands are illustrated in Table 2.3. The absorption bands that can be assigned to the individual functional groups, corresponding to the fundamental stretching and bending vibrations, lie above 1300 cm^{-1}. On the other hand, the sub-region between 1300-400 cm^{-1} exhibit vibrational bands overlapped that cannot be conclusively assigned to localized vibrations, although its feature/pattern is characteristic of each molecule, being this sub-region denominated fingerprint region.

Table 2.3. Typical IR absorption frequencies bands. Absorptions intensities are referred as strong (s), medium (m) and weak (w) (adapted from Pavia, 2004).

	Type of vibration		**Frequency (cm^{-1})**	**Intensity**
C-H	Alkanes	(stretch)	3000-2800	s
	$-CH_3$	(bend)	1450 and 1375	m
	$-CH_2-$	(bend)	1465	m
	Alkenes	(stretch)	3100-3000	m
		(out-of-plane bend)	1000-650	s
	Aromatics	(stretch)	3150-3050	s
		(out-of-plane bend)	900-690	s
	Alkyne	(stretch)	ca. 3300	s
	Aldehyde		2900-2800	w
			2800-2700	w
C-C	Alkane		Not interpretatively useful	
C=C	Alkene		1680-1600	m-w
	Aromatic		1600 and 1475	m-w
C≡C	Alkyne		2250-2100	m-w
C=O	Aldehyde		1740-1720	s
	Ketone		1725-1705	s
	Carboxylic acid		1725-1700	s
	Ester		1750-1730	s
	Amide		1680-1630	s
	Anhydride		1810 and 1760	s
	Acid chloride		1800	s
C−O	Alcohols, ethers, esters, carboxylic acids, anhydrides		1300-1000	s
O−H	Alcohols, phenols			
	Free		3650-3600	m
	H-bonded		3400-3200	m
	Carboxylic acids		3400-2400	m
N−H	Primary and secondary amines and amides			
	(stretch)		3500-3100	m
	(bend)		1640-1550	m-s
C−N	Amines		1350-1000	m-s
C=N	Imines and oxines		1690-1640	w-s
C≡N	Nitriles		2260-2240	m

In recent decades, FT-IR spectrometers are predominantly used to obtain IR spectra, with basis on the interference between two beams to yield an interferogram

(signal produced as a function of the change of pathlength between the two beams). The two domains of distance and frequency are interconvertible by the mathematical method of FT (Pavia *et al.*, 2009). The main component of the FT-IR instrument is the interferometer, the most commonly used the Michelson interferometer (Figure 2.13), consisting of two perpendicularly plane mirrors, one of which can travel in a direction perpendicular to the plane. A semi-reflecting film, the beam splitter, bisects the planes of these two mirrors, with the beam splitter material being chosen according to the region to be examined. For MIR and NIR regions, materials such as germanium or iron oxide are coated onto an 'infrared-transparent' substrate, such as potassium bromide. When a beam of IR radiation is passed into the beam splitter, half of the incident radiation will be reflected to one of the mirrors, whilst the other half will be transmitted to the other mirror. The two beams are reflected from those two mirrors, returning to the beam splitter where they recombine and interfere, with the pathlengths differences (differing wavelength contents) between the two beams causing both constructive and destructive interferences and originating the interferogram.

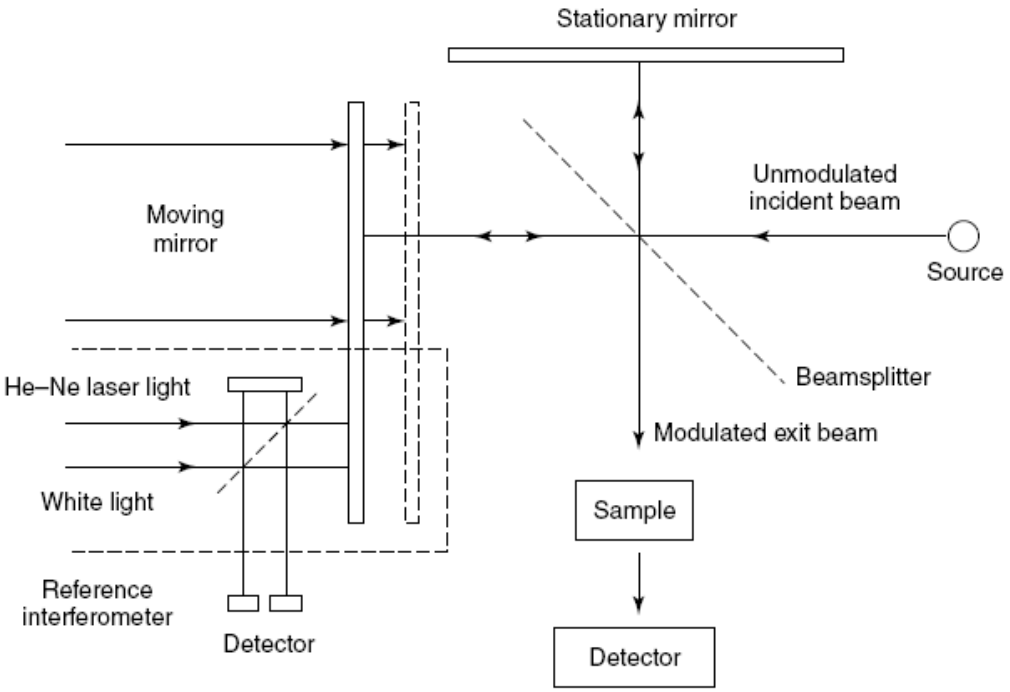

Figure 2.13. Schematic representation of a Michelson interferometer (reproduced from Stuart, 2004).

The interferogram is then oriented towards the sample and, as it passes through, the sample simultaneously absorbs all of the wavelengths that are found in its IR

spectrum. The modified interferogram signal, containing information about the amount of energy that was absorbed at every wavelength (frequency), reaches the detector where is compared to a reference laser beam resulting in a one time-domain signal. The mathematical process of FT is then implemented to extract the individual frequencies that were absorbed and to reconstruct the typical IR spectrum, with the relationship between the FTIR spectrum and the interferogram expressed in Eq. 2.15:

$$B(\upsilon) = \int_{-\infty}^{+\infty} I(x) \cos 2\pi\upsilon x.dx \qquad \text{[Eq.2.15]}$$

where B(υ) represents the intensity of the source as a function of frequency υ, i.e. the spectrum, I(*x*) is the intensity of the beam measured at the detector at a displacement of the movable mirror by *x* cm.

There are many sampling methods for measuring IR spectra of samples in different states of matter and for different applications. The attenuated total reflectance (ATR) method has, in recent years, revolutionized solid and liquid sample analyses because it minimizes the issue of sample preparation and spectral reproducibility, being the most widely used in the analysis of foodstuffs. This method consists in measuring the changes that occur in a totally internal reflected IR beam when it comes into contact with a sample, placed in a refractive index crystal (Figure 2.14). Briefly, an IR beam directed onto an optical dense crystal with a high refractive index (e.g. diamond, germanium, and zinc selenide), will undergo total internal reflection when the angle of incidence, θ, is greater than the critical angle (determined by the refractive index). The beam penetrates beyond the surface of the crystal into the sample and, in regions where the sample absorbs radiation selectively, the beam losses energy and its intensity is attenuated. The resultant attenuated radiation is measured at the end of the crystal and plotted as a function of wavelength by the spectrometer.

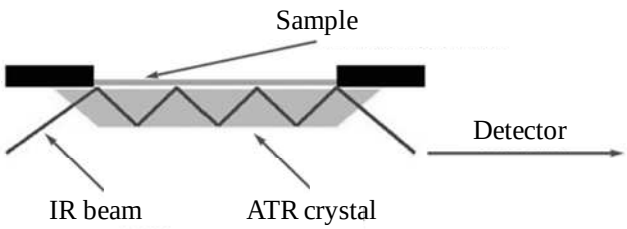

Figure 2.14 Schematic representation a multiple reflection attenuated total reflectance (ATR) (reproduced from: http://www.perkinelmer.com).

2.5 The metabonomics approach

2.5.1 Metabonomics: definition and applications in food science

The term metabonomics has been established to describe a system approach used to examine the changes in the low-molecular metabolites in a biofluid or an intact biological tissue. The definition of metabonomics was proposed in the 1990's, being expressed as "the quantitative measurement of the time-related multiparametric metabolic response of living systems to pathophysiological stimuli or genetic modification" (Nicholson *et al*., 1999). This approach has been extensively applied in the pharmaceutical and medical areas (Holmes *et al*., 2001; Garrod *et al*., 2001; Chen *et al.*, 2006; Graça *et al*., 2009; Jiang *et al*., 2010), particularly in the description of the overall organism response to stimuli such as disease, toxic agents or diet.

Metabonomic studies involve, primarily, the use of NMR spectroscopy and MS-based methods (the latter generally coupled with chromatographic methods: GC-MS and LC-MS) combined with multivariate analysis (MVA) to establish a metabolite profile of the sample and, therefore, a physiological state of the individual. This metabonomics approach has been extended to food science in two different areas: i) nutritional metabonomics, exploring the metabolic effects (by analysis of biological biofluid or tissue) of particular foods, food components or food combinations (i.e. to understand the global biochemical aspects of food when interacting with human organism) (Solanky *et al.*, 2005; Rezzi *et al*., 2007; Jiang *et al*., 2010); and ii) food metabonomics, defining the metabolic profile of a food product as its fingerprint and, thus using it as a reference of the state of the food (fresh vs. aged, adulterated, contamined, and genetically modified).

Food metabonomics has been extensively carried out, with particular emphasis on the detailed chemical characterization of foods, including milk (Ninonuevo *et al*., 2006; Agabriel *et al*., 2007), grapes and wine (Pereira *et al*., 2006; Rocha *et al*., 2006), tomato and tomato juice (Moco *et al*., 2006; Tiziani *et al*., 2006), fruit juices (Le Gall *et al*., 2001); vinegar (Caligiani *et al*., 2007) and beer (Duarte *et al*., 2002; Pinho *et al*., 2006); as well as on food quality and/or authenticity assessment. Authenticity studies have been particularly focused on the detection of adulterations and contaminations, such as in fruit juices (Gomez-Carracedo *et al*., 2004; Cuny *et al*., 2008; Spraul *et al*., 2009), olive oils (Vigli *et al*., 2003; James *et al*., 2004) and vinegars (Consonni and Gatti, 2004; Consonni *et al*., 2008); and in product origin classification, e.g. in wines (Ogrinc

et al., 2001; Brescia *et al*., 2003; Son *et al*., 2009), olive oils (Mannina *et al*., 2001; Mannina *et al*., 2003; D'Imperio *et al*., 2007) and fruit juices (Spraul *et al*., 2009). The monitoring of production processes or batch-to-batch reproducibility has also been studied, particularly in beer (Nord *et al*., 2004; Lachenmeier *et al*., 2005; Almeida *et al*., 2006), coffee (Charlton *et al*., 2002), wine (Kirwan *et al*., 2008), fruit juices (Reid *et al*., 2004; Spraul *et al*., 2009) and vinegars (Consonni and Gatti, 2004).

2.5.2 Analytical and data analysis tools

The main analytical techniques employed for metabonomics studies are based on NMR spectroscopy and MS-based methods (e.g. GC-MS and LC-MS), due to their great ability to provide compound structural information, with NMR enabling the detection of a wider group of compounds in a single run, whereas MS show higher sensitivity. The two analytical techniques are, therefore, complementary and their tandem use usually enables more detailed compound identification. In Table 2.1 are listed the main strengths and weakness of NMR and MS for metabolic profiling. Other analytical techniques, such as IR spectroscopy and arrayed electrochemical detection, have also been employed, however, due to limitations related with the reduced ability to obtain components structural information, the use of those techniques have been limited.

MVA of the collected data is of crucial importance for the outcome of metabonomic studies. In fact, aiming to interpret that profiles of metabolism, MVA have been used, usually involving the application of unsupervised methods, such as principal component analysis (PCA), hierarchical cluster analysis and non-linear mapping, and supervised methods, e.g. partial least squares (PLS), soft independent modeling of class analogy (SIMCA) and artificial neural networks (ANN) (Lindon and Nicholson, 2008).

Below, the main analytical techniques employed in metabonomics, i.e. NMR and MS, will be briefly described, as well as the data analysis methods applied in the work developed in this thesis.

Table 2.4. The relative strengths and weaknesses of NMR and MS for metabolic profiling (adapted from Lindon and Nicholson, 2008).

	NMR	MS
Detection limits	Low-micromolar at typical observation frequencies (600 MHz), but nanomolar using cryoprobes	Picomolar with standard techniques, but can be much lower with special techniques
Universality of metabolite detection	If metabolite contains hydrogens it will be detected, assuming the concentration is sufficient or protein binning does not cause marked line broadening	Usually needs a more targeted approach. There can be problems with poor chromatographic separation; with the loss of metabolites in void volumes; with ion suppression (but this is reduced when using UPLC); lack of ionization; ability to run both +eV and –eV ion detection gives extra information
Sample handling	Whole sample analyzed in one measurement	Different LC packings and conditions for different classes of metabolite; usually samples have to be extracted into a suitable solvent; samples have to be aliquoted but some recent studies have avoided the need for chromatography
Amount of sample used	Typically 200-400 μL, but much less for microcoil probes, down to 5 μL	Low μL range
Sample recovery	Technique is nondestructive	Technique is destructive but only small amounts used
Analytical reproducibility	Very high	Fair
Sample preparation	Minimal: addition of buffer, D_2O and chemical shift reference (not always required)	Can be substantial; often needs different LC columns and proteins precipitation
Ease of molecular identification	High, both from database of authentic material and by-self-consistent analysis of 1D and 2D spectra	Difficult, often only the molecular ion is available; this needs extra experiments, such as routine tandem MS; GC-MS is generally better with accurate retention times and comprehensive databases of spectra
Time to collect basic data	5 min for 1D ^{1}H NMR	10 min for UPLC-MS run
Quantitation	1-5%	5% intraday and interday is now common with or without prior chromatography
Robustness of instruments	High	Low
Molecular dynamics information	Yes, from T_1 and T_2 relaxation time and diffusion coefficient measurements	No
Analysis of tissues samples	Yes, using MAS NMR	No
Availability of databases	Not yet comprehensive but increasing; several are available freely on the web; some commercial products also exists	Comprehensive databases for electron impact MS allow spectral comparison; for electrospray ionization, as is usual in LC-MS, only mass values can be compared

2.5.2.1 Analytical techniques

NMR spectroscopy

Although lesser sensitive than MS methods (low micromolar vs. picomolar from MS), NMR has the advantage of allowing detecting a wide range of families of compounds in a single run, with higher reproducibility. Commonly, metabonomics studies are based on 1D ^{1}H NMR spectra data, which can be recorded in only a few minutes. Fully automated high-throughput NMR systems have been introduced (Spraul *et al.*, 2009), improving the NMR potential for routine application in industry. Essentially, metabolites assignment is performed with basis in comparing chemical shifts, multiplicity and coupling constants data. For complex mixtures, signal assignment is hindered by major spectral overlap in ^{1}H NMR spectra. Thus, 2D NMR methods are applied to help identify metabolites, increasing dispersion (information is spread in two dimensions) and elucidating the connectivities between the signals. 2D NMR experiments applied include correlation spectroscopy (COSY), TOCSY, *J-resolved* and heteronuclear shift correlations experiments. In recent years, a large volume of NMR spectral databases has become available, either compiled in literature (Nicholson *et al.*, 1995, Fan, 1996; Fan 2008) and from commercial sources (http://acdlabs.com, www.bruker-biospin.com), as well as freely available in public websites (e.g. http://www.hmdb.ca/, http://riodb01.ibase.aist.go.jp/).

MS-based methods

MS-based methods have been widely used in metabolic fingerprinting and metabolites identification, mainly combined with chromatographic (GC, LC, or, more recently, ultra performance LC (UPLC) methods) and capillary electrophoresis methods. Furthermore, sample preparation can become time-consuming since, in most studies, an initial sample extraction is required. The MS technique is inherently more sensitive than NMR, however, it is less reproducible and robust and requires different samples preparation and separation techniques depending on the analytes. Quantification can be impaired by variable ionization and ion suppression effects, showing a consequent a higher error percentage. A great variety of software tools are available for structural elucidation, such as MassFrontier, ACD/MS Manager, NIST MS Search and Sierra's APEX software (Kind and Fiehn, 2010), as well as large public compounds databases (already referred in NMR spectroscopy section).

2.5.2.2 Multivariate analysis (MVA)

Prior to MVA, the used data matrices must be pre-treated for multivariate purposes, in order to avoid possible artefacts which could potentially introduce some systematic and random variations in the data matrices. The steps involved in the data pre-treatment typically consisted at a minimum of: i) data processing, ii) data normalization and iii) data scaling.

Processing

Peak shift is a major problem in metabonomics, being attributed mainly to instrumental drift, random variation or chemistry of sample (e.g. solvent, pH, interactions with stationary phase and ion strength), the latter being considered the largest source of peak shift (Aberg *et al.*, 2009). Two main methods have been used to overcome the problem, namely spectral binning and peak alignment. Spectral binning (mainly used in NMR data), consists in subdividing the spectra into a number of regions along the chemical shift axis, and the total area within each bin being summed to provide an integral so that the intensities of the peaks in such defined spectral regions are extracted. The size of the bins can be of a fixed width, variably sized using manual inspection or automated procedures (Craig *et al.*, 2006; Weljie *et al.*, 2006), with the accepted criteria for binning width of 0.01-0.04 ppm (Cloarec *et al.*, 2005). Peak alignment is an alternative to spectral binning, being used in both spectroscopic and chromatographic methods. Most alignment methods work in pair-wise fashion by aligning either only pairs of samples or multiple samples against a selected reference sample or template. Several alignment methods have been employed in metabonomics including correlation optimized warping (COW), dynamic time warping (DTW), parametric time warping (PTW), Fast Fourier Transform (FFT), and XCMS algorithm (Katajamaa and Oresic, 2007; Aberg *et al.*, 2009).

Normalization

Sample (spectrum in the case of spectroscopic and run for chromatographic methods) normalization is required to reduce or minimize the effects of systematic variations of samples overall concentration. One of the most commonly applied method of normalization is the normalization to total area, which involves setting each observation to have unit total intensity by expressing each data point as a fraction of the total observation integral (Craig *et al.*, 2006; Lindon *et al.*, 2007).

Data scaling

The optional final pre-treatment step is the data scaling, which is a column operation (i.e., on each variable across all samples), being aimed to the weighting of data variables. A number of such scaling methods are commonly used, namely mean-centering, unit-variance, Pareto and logarithm. In those scaling methods, each column can be given a mean of zero by subtracting the column mean from each value in the column (mean-centering method). Secondly, each column can be scaled so that it has: unit variance (by dividing each variable in the column by the standard deviation of the column and, therefore, giving equal weight to each variable), Pareto (each variable is divided by the square root of the standard deviation of the column values) or logarithmic scaling.

In this thesis, the MVA methods of unsupervised method of PCA, supervised methods of PLS regression and PLS - discriminant analysis (PLS-DA), data fusion methodology of outer-product analysis (OPA), the two-dimensional (2D) correlation analysis and statistical heterospectroscopy (SHY) were employed. Their basic principles are described below.

Principal component analysis (PCA)

PCA is essentially a descriptive method often used for data reduction and exploratory analysis on high-dimensional data sets. It helps establishing the presence of any intrinsic class-related patterns or clusters in the data set without any inclusion of information concerning the samples classification , i.e. it is an unsupervised method (Jollife, 1986), as well as enables the detection of outliers in the data sets.

Considering an analytical data set assembled into a matrix **X**, the PCA decomposes that original matrix into two matrices, the loadings, **P**, and the scores, **T**. This decomposition is expressed in Eq. 2.16, where, additionally, **E**, holds residual variation.

$$\mathbf{X} = \mathbf{T}\,\mathbf{P}^{\mathbf{T}} + \mathbf{E} \qquad \text{[Eq. 2.16]}$$

The principal components, PC's, are linear combinations of the original variables, calculated in order to maximize the variance of the objects, with the successive PC's accounting for decreasing amount of explained variance. The scores matrix represents the new coordinates of each object, i.e. the relationships between the samples, when

projected in the new axes. On the other hand, the relationship between new and old axes, i.e. the importance of each variable to the scores distribution, is expressed in the loadings matrix explained by each PC's.

Partial least squares (PLS)

Regarding PLS regression method, its purpose is to build a linear regression model that enables prediction of a desired characteristic from a measured multivariate signal (e.g. spectra), being known as supervised method (Wold *et al.*, 2001). The PLS is a predictive two-block regression method based on estimated latent variables and applies to the simultaneous analysis of two data sets for the same object: matrix **X**, comprising analytical data of a set of samples; and matrix **Y**, a quantitative response matrix. PLS regression model can be expressed by:

$$\mathbf{X} = \mathbf{T}\,\mathbf{P}^{\mathbf{T}} + \mathbf{E} \qquad \text{[Eq. 2.17]}$$

$$\mathbf{Y} = \mathbf{T}\,\mathbf{C}^{\mathbf{T}} + \mathbf{F} \qquad \text{[Eq. 2.18]}$$

where, **T** is a matrix of PLS scores common to **X** and **Y**, **P** and **C** are the matrices of X- and Y-loadings, respectively, used to build the models, and **E** and **F** are the X- and Y-residuals representing the part of the data not explained by the model. The scores matrix **T** is calculated as the projection of the data **X** onto a weights matrix **W**:

$$\mathbf{T} = \mathbf{X}\,\mathbf{W} \qquad \text{[Eq. 2.19]}$$

Unlike PCA, the X- and Y-loadings in PLS, **P** and **C**, are no longer orthogonal due to the constraints in finding common scores **T** in the decomposition. Using this, the equation [Eqs. 2.17 - 2.18] can be rearranged giving the following PLS solution to the regression:

$$\mathbf{Y} = \mathbf{X}\,\mathbf{B} + \mathbf{F} \qquad \text{[Eq. 2.20]}$$

where **B** refers to the regression coefficients matrix. Furthermore, in PLS regression when the response variable, **Y**, is a categorical, expressing the class membership of the statistical units (Viereck *et al.*, 2008), the method is called Partial Least-Squares Discriminant Analysis (PLS-DA).

Regarding PLS validation, a commonly used parameter to assess the prediction ability in PLS regression models is the root mean squares error of prediction (RMSEP) (Faber and Rajko, 2007), which is computed as:

$$RMSEP = \sqrt{(1/n)\sum_i (y_i - \hat{y}_i)^2} \qquad \text{[Eq 2.21]}$$

where $\hat{y}_i$ and y_i represent the predicted and the measured response value of the *i*th $_{sample}$, respectively, and *n* refers to the number of prediction samples. The lower the RMSEP the higher would be the prediction ability of the model. Furthermore, ideally, the use of the optimum number of components minimizes the RMSEP values of the model.

To test the validity of PLS models against over-fitting, representative of the good predictive/classificatory ability of the model, several approaches can be used, both internal and external validation. The former approach includes the methods of cross-validation and the Monte Carlo cross validation (MCCV). Various parameters have been used to evaluate models classification, e.g. Q^2 values, number of misclassifications, many combinations of sums or ratios of True Positives, True Negatives and False Positives and False Negatives of a confusion matrix and by analysis of receiver operating characteristic (ROC) plots (Westerhuis *et al.*, 2008). Particularly in this thesis, the cross-validated Q^2 value parameter was used to validate the build models, being defined as follows:

$$Q^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2} \qquad \text{[Eq.2.22]}$$

where $\bar{y}$ refers to the mean value of all samples. The optimal Q^2 value would be of 1, however that required that the class prediction of each sample should be exactly equal to its class label. The MCCV method is based on the Monte Carlo framework, which iteratively performs the cross validation N times (generally higher than 1000 iterations). At each time, the test data and the training data are chosen randomly.

Outer product analysis (OPA)

Outer product analysis (OPA) is a method of data fusion which combines two (or even more) data sets by first calculating an outer product matrix from each pair of samples vectors (Barros *et al.*, 1997; Jaillais *et al.*, 2005; Rutledge and Bouveresse, 2007; Barros *et al.*, 2008). Each matrix is then unfolded to produce a very wide vector containing all possible product combinations between the two data sets. It is then possible to perform statistical analyses, such as PCA, PLS regression or Factorial Discriminant Analysis (FDA), on the resulting outer product matrix in order to highlight relations between the two data sets, both linear and non-linear relations. This operation will, for instance, emphasize co-evolutions of spectral regions in signals acquired in two different domains (heterospectral) or within the same domain (homospectral). This data

fusion technique has been applied in different contexts of instrumental analysis and in data sets from numerous analytical techniques, such as MIR and NIR, Raman and NMR. Example of OPA methodology is expressed in Figure 2.15.

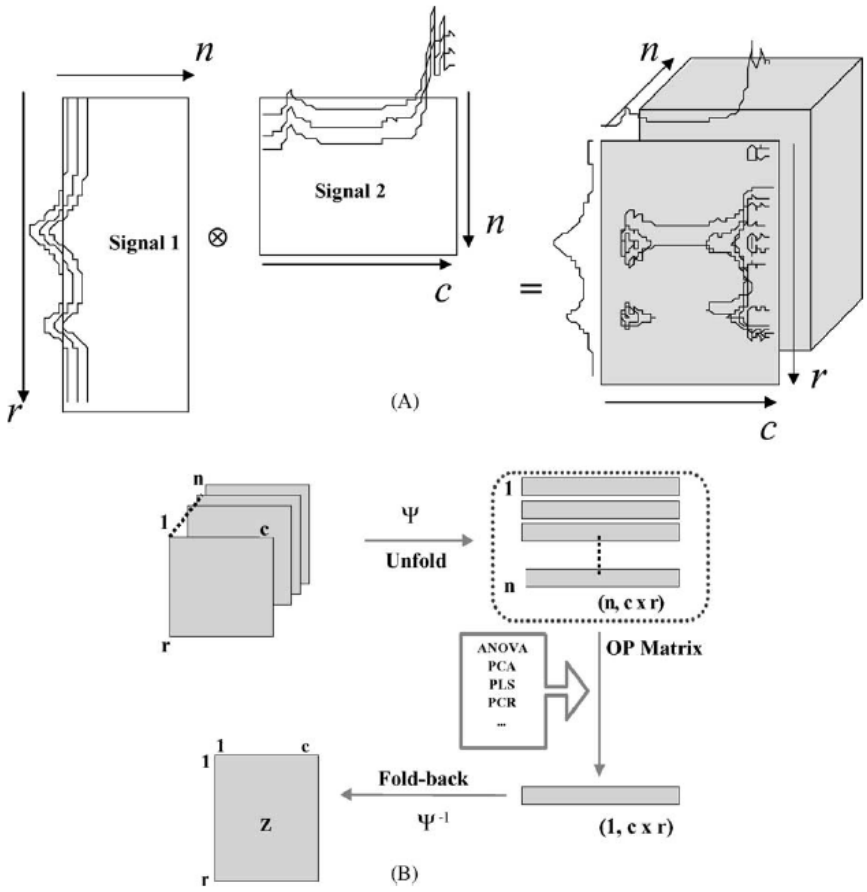

Figure 2.15. OPA procedure for two signals acquired for each sample: (A) unfolding of the OP matrices, concatenation of the vectors, statistical analysis of the resulting OP matrix and (B) refolding of the vectors of calculated values (reproduced from Jaillais *et al.*, 2005).

Essentially, the OPA procedure starts by calculating the outer products of the two vectors corresponding to the signals acquired for each of the *n* samples in the two domains. All the intensities of one domain are multiplied by all intensities in the other, resulting in a set of *n* data matrices containing all possible combinations of intensities in the two domains. This outer product (OP) of pairs of signals vectors of lengths *r* and *c* for the *n* samples gives n (*r* rows, *c* columns) matrices (Figure 2.15A). These matrices are then unfolded to give *n* (1, *r* x *c*) row vectors. This procedure corresponds to a mutual weighting of each signal by the other:

i) if the intensities are simultaneously high in the two domains, the product is higher;
ii) if the intensities are simultaneously low in the two domains, the product is lower;
iii) if one of the intensities is high and the other low, the resulting product tends to an intermediate value.

By concatenating row-wise all row vectors, a matrix (*n*, *r* x *c*) is produced, termed the OP matrix (Figure 2.17B).The resulting OP matrix can be analyzed by several multivariate analysis methods (e.g. PCA, PLS or FDA), after which, each latent vector produced, e.g. loadings or B-coefficients, can be folded back to give a latent **Z** matrix of dimensions (*r* x *c*), which may be examined to detected the relations between the two domains, based on the properties of the applied chemometric technique (Barros *et al.*, 2008). In some cases, the resulting OP matrix is very wide and demanding in terms of computer resources (time and memory), being used the method of principal component transform (PCT) in order to overcome this problem, allowing working in a reduced, compressed space (Barros *et al*., 2004; Barros *et al*., 2007; Barros *et al*., 2008).

Correlation methods: two dimensional (2D) correlation analysis and statistical heterospectroscopy (SHY)

Two dimensional (2D) correlation spectroscopy (in this thesis mentioned as 2D correlation analysis in order distinguish from the NMR terminology for the 2D correlation spectroscopy experiment, COSY) was proposed as an extension of original 2D correlation analysis, allowing producing 2D correlation spectra from systematic variations of the spectra having an arbitrary and complex wave form (Noda *et al.*, 1993). The 2D correlation analysis enables to monitor spectral intensity fluctuation as a function of not only time-related but also of other physical variables, e.g. temperature, pressure, distance or even concentrations.

This technique can be described as the quantitative comparison of the features of spectral intensity variations along an external perturbation, observed at two different spectral variables and over a finite observation interval, considering not the specific individual spectra themselves but the dynamics and changes associated with spectral features. This method considers not the specific individual spectra themselves, but the dynamics and changes associated with the spectral features, being those changes visualized and investigated by examination of spectral maps. In 2D correlation analysis,

the correlation between signals may be treated as a complex number comprising two orthogonal components known as synchronous, Φ, and asynchronous, Ψ, correlation intensity (Noda, 2004). These components are defined by:

$$\Phi = \bar{\mathbf{A}}^T \bar{\mathbf{A}} \qquad \text{[Eq. 2.23]}$$

$$\Psi = \bar{\mathbf{A}}^T N \bar{\mathbf{A}} \qquad \text{[Eq. 2.24]}$$

corresponds to the matrix of corrected dynamic spectra, recorded at $i = 1$ to n sequential time intervals using $j = 1$ to m spectral variables (chemical shifts values), and N is the Hilbert-Noda matrix (size m x m) that serves to extract out of phase portions of the signals: if $i = k$, $N_{i,k} = 0$, otherwise $N_{i,k} = 1/\pi\,(i\text{-}k)$ (Noda, 2004; Kirwan *et al.*, 2009). The synchronous matrix, Φ, represents simultaneous or coincidental changes, expressed as positive (in the same direction) or negative (in opposite direction) correlations between the spectral data points, whereas asynchronous matrix, Ψ, represents the sequential (or "*out of phase*") changes in the spectral intensities.

These two matrices, synchronous and asynchronous, can be investigated by examination of their corresponding spectral maps. In the synchronous maps, the interpretation of the pairs of variables arising is direct, as positive correlations represent signals varying in the same direction, whereas negative relates signals varying in opposite direction. On the other hand, the interpretation of the asynchronous data relies on applying 'Noda rule'. Briefly, for pairs of variables (data points) positively correlated in the synchronous matrix, the sign of an asynchronous peak is positive if the intensity change at a variable occurs before (or at a higher extent) that of a second variable in the sequential order, whereas the sign becomes negative if the change at the first variable occurs latter (or to a lesser extent). This sign rule is reversed if the synchronous correlation intensity at the same coordinate is negative.

Statistical heterospectroscopy (SHY) is a statistical strategy for the co-analysis of multiple spectroscopic data sets acquired in parallel on the same samples (Crockford *et al.*, 2008). This approach, which evolved from the similar concept of statistical total correlation spectroscopy (STOCSY) applied to NMR spectra, operates through the analysis of the intrinsic covariance between signals intensities obtained from NMR spectroscopy and from different spectroscopic sources, mainly MS-based methods (Crockford *et al.*, 2006; Crockford *et al.*, 2008; Maher *et al.*, 2011) across cohorts of

samples, showing great potential as an analytical tool for enhancing feature extraction from different data types from the same samples.

SHY is based on the determination of linear relations (correlations) between signal intensities for each variable of two domains, X and Y, enabling the identification of highly correlated pairs of variables. SHY consists in a previous autoscaling of each matrix and a following construction of the correlation matrix **C** according to:

$$\mathbf{C} = [1/(n-1)]\ \mathbf{X_A}^T\ \mathbf{Y_A} \qquad \text{[Eq. 2.25]}$$

where $\mathbf{X_A}$ and $\mathbf{Y_A}$ refer to the autoscaled matrices **X** and **Y**, respectively, and n is the number of analyses per data set. The detected highly correlated pairs of variables can correspond to the identification of variables from the same molecule (if detected in both domains), metabolites in the same metabolic pathway whose concentrations are inter-dependent or under common regulatory mechanism (Cloarec *et al.*, 2005).

3. BEER COMPOSITION AND COMPOSITIONAL VARIABILITY

3.1 Introduction

The invaluable ability of NMR spectroscopy to provide compound structural information has been extensively exploited in the identification and characterization of a wide range of beer components. In fact, extensive work has been performed in the chemical characterization of beer by 1D and 2D NMR and hyphenated NMR (LC-NMR and LC-NMR/MS) enabling a considerable database of compounds to be established (Duarte *et al.*, 2003; Gil *et al.*, 2003; Nord et *al.*, 2004; Almeida *et al.*, 2006). In this chapter, following those same methodologies, the detailed chemical characterization of a lager beer was assessed to attempt improving beer characterization. The fact that a new samples preparation step was employed (pH adjustment to ~1.90), compared to previous studies, makes this initial characterization extremely important for the following metabonomic studies performed in this thesis.

In the second part of this chapter, the application of NMR/MVA methodology for evaluating beer variability aiming the short-term monitoring of beers is described. This issue was addressed in previous reports, where beer (from the same brand) variability between different brewing sites and dates, however located at different countries was studied (Almeida *et al.*, 2006). In that study, a clear separation of beers differing in brewing sites was noted, suggesting small differences in yeast generation and/or characteristics of brewing steps (malting and mashing steps). Besides this, the fact that water origin is distinct in different countries (contrary to plant origin which is kept the same) may play an important role. In this thesis, since the studied beers were produced nationally, a better control of the starting materials characteristics was ensured which, together with the use of more uniform brewing protocols, may lead to an overall finer control of brewing process and of beer composition. In this way, PCA and PLS-DA methods were applied to ^{1}H NMR data corresponding to a set of lager beers from the same brand, produced in three different brewing sites (located in the same country, Portugal) and three different dates in order to assess finer compositional changes occurring in beers, produced under well-controlled and uniform industrial environment, in order to achieve an overall finer control of the process and of the final composition of beer.

3.2 Experimental section

3.2.1 Beer samples and sample preparation

All lager beer samples were of the same brand and kindly donated by UNICER, Bebidas de Portugal. A total of 35 beers were studied (Table 3.1) originating from three different brewing sites (**A**, **B** and **C**) located in Portugal, and three production dates (date **a**, **b**, **c**), corresponding to maximum differences of 1 month between dates. For each site and date 4 bottles were considered in order to evaluate batch reproducibility (except for site **C** and date **c** for which only 3 samples were analyzed, due to breakage of the 4th bottle). The bottled samples were kept unopened at 4^{o}C, for up to 15 days prior to analysis.

Table 3.1. Beer sample sets employed for the compositional variability studies.

Beer sample groups	Production date	No. bottles/date	Bottles pH
	Site A		
Date **a** (code **Aa**)	March 2006	4	4.49
Date **b** (code **Ab**)	April 2006	4	4.36
Date **c** (code **Ac**)	April 2006	4	4.37
	Site B		
Date **a** (code **Ba**)	March 2006	4	4.41
Date **b** (code **Bb**)	March 2006	4	4.32
Date **c** (code **Bc**)	April 2006	4	4.29
	Site C		
Date **a** (code **Ca**)	February 2006	3	4.54
Date **b** (code **Cb**)	March 2006	4	4.43
Date **c** (code **Cc**)	March 2006	4	4.38

Beer samples (10 mL) were degassed in an ultrasonic bath for 10 minutes and prepared to contain 10% of D_2O and 0.025% 3-(trimethylsilyl)propionic-2,2,3,3-*d4* acid, sodium salt (TSP-*d4*) as chemical shift and intensity reference. Sample pH was adjusted to 1.90 ± 0.05 by adding 8-10 µL HCl 5% in D_2O (to ensure extensive protonation of carboxylic groups and reduce peak shifting), prior to sample transfer into 5 mm NMR tubes.

For the study of the effect of pH on the ^{1}H NMR spectra of beer, a beer sample was degassed in an ultrasonic bath for 10 minutes and four aliquots were prepared to contain 10% of D_2O and 0.025% TSP-*d4*, and therein pH was adjusted to i) 4.00, ii) 3.0, iii) 2.50, and iv) 1.90 (adding HCl 5% in D_2O). For specific studies on pyruvic acid, a standard solution of that acid (concentration of 0.57 mmol L^{-1}) was divided into five

aliquots and, after addition of 10% of D_2O and 0.025% TSP-*d4*, each aliquot pH was adjusted to i) 4.00, ii) 3.5, iii) 3.0, and iv) 2.5, and v) 1.9.

To study beer composition as a function of storage conditions, 9 aliquots from the same beer bottle were transferred into 5 mL vials and kept at i) -20ºC and ii) -70 ºC, during 90 days to assess its compositional stability along time of storage. At days 15, 30, 60 and 90, stored samples were thawed, degassed (10 minutes, ultrasound), prepared to contain 10% of D_2O and 0.025% TSP-*d4*, and sample pH was adjusted to 1.90 ± 0.05 as described above.

3.2.2 NMR measurements and signal integration

All NMR spectra were recorded on a Bruker Avance DRX 500 spectrometer (Bruker, BioSpin, Rheinstetten, Germany), operating at 500.13 MHz for proton and 125.77 MHz for carbon, equipped with an actively shielded gradient unit with a maximum gradient strength output of 53.5 G/ cm and a 5 mm inverse probe, at 300 K.

1D ^{1}H NMR: for each beer sample, two ^{1}H NMR spectra (replicates) were acquired using the *noesypr1d* pulse sequence (Bruker pulse program library) with water presaturation. 128 transients were collected into 32768 (32 K) data points, a spectral width of 8013 Hz, acquisition time of 2.0 s and relaxation delay of 5 s. Each FID was zero-filled to 64 k points and multiplied by a 0.3 Hz exponential line-broadening function prior to Fourier transform (FT). 1D spectra of standard solutions were also acquired as described above. All spectra were manually baseline corrected and chemical shifts referenced internally to the TSP peak.

2D NMR: the total correlation (TOCSY) spectra of selected beer samples were acquired using the pulse program *lcmlevpcpstp*, in the phase sensitive mode using time proportional phase increment (TPPI), with a MLEV17 spin lock pulse sequence and presaturation of water and ethanol resonances. 1024 data points with 56 transients per increment and 512 increments were collected using a spectral width of 5482 Hz in both dimensions, a mixing time of 80 ms and a relaxation delay of 1.5 seconds. The data was zero-filled to 1024 data points in the f_1 dimension and a sine-bell function was applied prior to FT.

The ^{1}H-^{13}C phase sensitive (echo/antiecho) heteronuclear single quantum correlation (HSQC) spectra were recorded with the inverse detection and ^{13}C decoupling during acquisition, using the pulse program *invietgpsi*. 2048 data points with

80 scans per increment and 300 increments were acquired, with spectral width of 5482 Hz and 25154 Hz in ^{1}H and ^{13}C dimensions, respectively, a refocusing delay equal to 1/4 $J_{C\text{-}H}$ (1.74 ms) and a relaxation delay between pulses of 1.5 s. The FIDs were weighted using a sine-bell function in both dimensions and zero-filled in the f_1 dimension to 512 data points.

2D homonuclear *J-resolved* spectra using the *lcjrespr* pulse program, with shaped pulse for water and ethanol suppression (triple suppression), were acquired. 16384 data points were collected with 16 transients per each of 64 increments, using a spectral width of 8012 Hz and 16 Hz in *f2* and *f1* dimensions, respectively. *J-resolved* spectra were also acquired in Bruker BioSpin, Germany, in a 400 MHZ spectrometer, using the *jresgpprqf* pulse program, with gradient pulse for only water suppression. 8192 data points were collected with 8 transients per each of 40 increments, using a spectral width of 6400 Hz and 78 Hz in *f2* and *f1* dimensions (*J* dimension), respectively. In both experiments, the FIDs were weighted in both dimensions by a sine-bell and zero-filled in the f_1 dimension to 256 data points, prior to FT.

LC-NMR and LC-MS: the instrumentation used consisted of a HLPC system with an Agilent HP 1200 solvent delivery system with a quaternary pump, solvent degasser, autosampler, diode array detector (DAD) and a column oven (Agilent, Waldbronn, Germany). The Agilent HP 1200 system was connected through a BPSU-36/2 interface to an Avance 600 MHZ spectrometer with cryoprobe (Bruker, BioSpin, Rheinstetten, Germany) and a Esquire 3000 electrospray ion trap mass spectrometer (Bruker Daltonics, Bremen, Germany). For beer analysis, chromatographic separation were performed by injection of beer samples onto a Phenomenex RP-C18 column at 35ºC, using a mobile phase (pH 2.0) consisting of an mixture of solution A (deuterated water + 0.1% formic acid deuterated to 90%) and solution B (acetonitrile + 0.1% formic acid deuterated to 90%), with the following gradient program: 3% to 40 % of solution B from 0 to 60 mins and then from 40 to 90% of B from 60 to 70 mins; and with DAD detection set at 280 nm and 200-300 nm.

For LC-NMR analysis, 100 μL injection volumes of beer sample were used. ^{1}H NMR spectra were recorded at 600.13 MHz, as mentioned before, with a cryoprobe and 120 μL active volume), using the time-sliced stopflow method. Briefly, time slice method consists in stopping the flow at short, constant time intervals, independently of LC peak positions. This allows a pseudo on-flow experiment to be obtained with

reduced time resolution along the chromatographic axis and improved NMR sensitivity due to the longer acquisition time. The *lc1pnf2* pulse sequence from the Bruker pulse program library was used, with water presaturation at 2817.22 Hz, a flow of 0.3 mL/min, and a total run time of 70 minutes. Every 30 s, the flow was stopped and a 1D ^{1}H NMR spectrum collected with 64 transients, 32k data points, spectral width of 12019 Hz, and 1.36 s of acquisition time. Each FID was multiplied by an exponential function (1 Hz line broadening), manually phased and baseline corrected. When using the *time-slice* method, no MS data can be acquired due to the interruption of the flow during the NMR measurements. To overcome this, LC-MS analysis was performed during the chromatographic run. The injected volume was 20 μL, and mass spectra were collected in positive mode, in the mass range of 50-1000 *m/z*, with auto MS. A spray voltage of 3500 V was applied, N_2 was used as sheath gas at 60 psi, 11 L/min and at 350 °C as dry gas.

All NMR peak assignments were carried based on previous work (Duarte *et al.*, 2002; Gil *et al.*, 2003; Duarte *et al.*, 2003; Nord *et al.*, 2004; Almeida *et al.*, 2006), data from 1D and 2D NMR and LC-NMR experiments of beer and standard solutions of relevant compounds at pH 1.90 and data from both Bruker Biorefcode spectral base (at pH 3.0) as well as other public databases, including Spectral Database for Organic Compounds – SDBS (organized by National Institute Industrial Science and Technology) and Human Metabolome Database (supported by Departments of Computing Science and Biological Sciences, university of Alberta).

Peak integrations were carried out with AMIX.3.9.5, Bruker BioSpin and peak areas were normalized relatively to total spectral area. For evaluation of dextrin changes, integration of specific anomeric signals was performed according to previous reports (Jodelet *et al.*, 1998; Duarte *et al.*, 2003): a) the “average number of glucose units per molecule” was calculated through the ratio of the [summed areas of α(1→4) and α (1→6) H1 protons, at 5.37 and 4.95 ppm respectively] and the [summed areas of reducing H1 protons (α and β anomers) at 5.21 and 4.64 ppm] and b) the “average number of branching points per molecule” was estimated by the ratio of the integral of α (1→6) H1 protons and that of α (1→4) H1 protons, multiplied by the total number of glycosidic linkages (given by the average number of glucose units per molecule but one).

3.2.3 Multivariate analysis

For PCA and PLS-DA models, each spectrum was divided into variable width buckets, using Amix-Viewer 3.9.5 software. The water (4.7-4.9 ppm) and ethanol (1.0-1.3 ppm and 3.6-3.7 ppm) regions were excluded from the matrix, as well as the 0.0-0.8 ppm region where no sample signals resonate. The spectral regions considered for multivariate analysis were either the total spectral window (0.8-10.0 ppm) or the aliphatic (0.8-3.1 ppm), anomeric (4.5-5.8 ppm) and aromatic (5.8-10.0 ppm) regions separately. Each spectrum was normalized by dividing each variable by total spectral area and each variable was i) mean centered; or ii) autoscaled, using SIMCA-P 11.5 software.

3.3 Study of the effects of sample pH and storage

As beer pH, usually, varies between 4.0 to 4.5, the first approach was to analyze beer without any pH adjustment (pH ~ 4.20). However, problems in terms of significant signals shift due to small pH variations (± 0.1) between beer samples (particularly for histidine, acetic and citric acids) and difficulties in the assignment of some metabolites, such as GABA and malic acid, were observed. To overcome this, pH was adjusted to 1.90 ± 0.03. Figure 3.1 shows the ^{1}H NMR spectra of beer with different pH values: i) 4.0, ii) 3.0, iii) 2.5, and iv) 1.9. In the experiments performed, it was seen that, when using more acidic pH values, peak shifts are reduced due to an extensive protonation of carboxylic groups thus minimizing peak shifts, especially for the signals arising from histidine (C2H, ring doublet, at 8.66 ppm), proline (γ CH_2, β'CH and β'CH multiplets, at 2.02, 2.11 and 2.37 ppm, respectively) and acetic (CH_3 singlet, at 2.08 ppm), citric (α, β CH_2 and α', β' CH_2 doublets, at 2.83 and 3.00 ppm, respectively) and succinic (α, β CH_2 singlet, at 2.65 ppm) acids. Furthermore, working at pH 1.9, signals overlap has decreased, particularly in the aliphatic region. Indeed, as shown in Figure 3.1 oval shapes 1 and 2, for GABA (α CH_2 triplet, at 2.50 ppm) and malic acid (β' CH and β CH doublet of doublets, at 2.89 and 2.92 ppm, respectively), peaks overlap decreased, enabling their improved identification in the ^{1}H NMR spectra and signal integration. A better separation was also observed between the signals in the region (8.0-8.5 ppm), referring to adenosine and/or inosine (C8H, ring and C2H, ring singlets, at 8.43 and 8.54 ppm, respectively), cytidine (C5H, ring doublet, at 8.12 ppm) and formic acid (HCOOH singlet, at 8.23 ppm) (not shown).

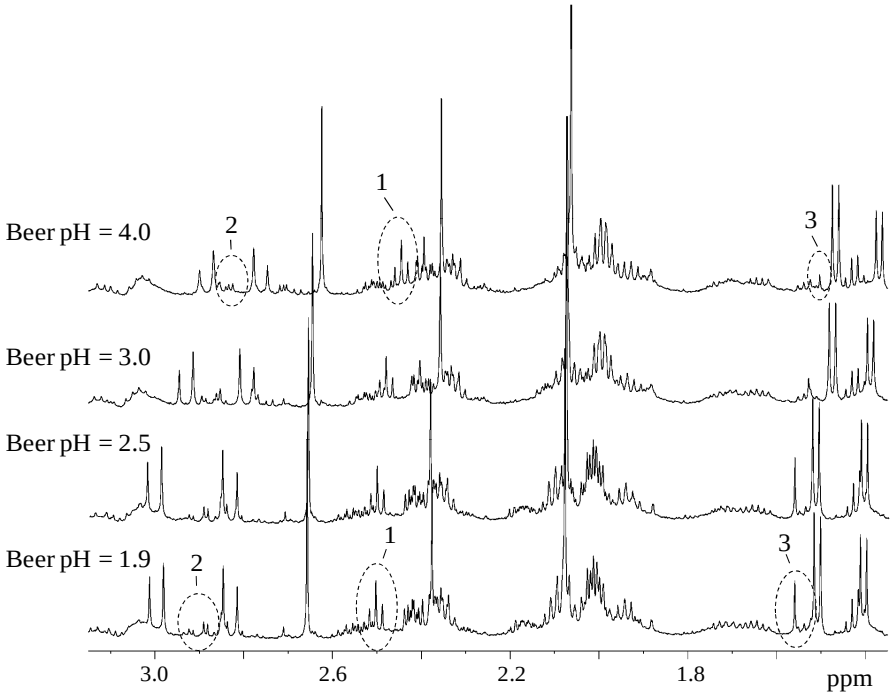

Figure 3.1. Inset of the aliphatic region (1.4-3.1 ppm) of 500 MHz ^{1}H NMR spectra of a lager beer sample, at different pH values: 4.0, 3.0, 2.5 and 1.9. Major variations between pH at 4.0 and at 1.9 are highlighted by oval shapes: 1), γ-aminobutyric acid (GABA); 2), malic acid; and 3), pyruvic and its hydrate form.

At acidic pH, pyruvic acid is converted to its hydrate form. In fact, pyruvate ($C_3H_4O_3$), in aqueous solutions, reversibly undergoes hydration into its hydrate form ($C_3H_6O_4$), with the hydrated species concentration increasing at more acidic pH and lower temperatures (Meany, 2007). Consequently, the low intensity signal corresponding to the hydrate form (singlet, at 1.50 ppm) in the spectrum at pH = 4.0 is replaced by a much higher intensity signal located at lower field (singlet, at 1.56 ppm) at pH = 1.9 (Figure 3.1, oval shape 3). This observation was confirmed by performing o standards solutions of pyruvic acid (concentration of 0.57 mmol L^{-1}) at different pH values (pH values: 4.0, 3.5, 3.0, 2.5 and 1.9) (not shown).

Beer stability of beer samples was assessed by visual comparison of the spectra obtained for beer aliquots stored during 90 days at i) -20 ºC and ii) -70 ºC (Figure 3.2). A similar profile is seen, even after 90 days of storage and independently of storage temperature. The only differences noted were the signals of pyruvic acid (at 2.40 ppm) and an unassigned signal (at 8.58 ppm), due to slight differences in the samples pH after preparation. Signal integration was also performed for acetic, citric, gallic, lactic, malic,

pyruvic and succinic acids, higher alcohols, alanine, GABA, histidine, phenylalanine, proline, tyrosine, tyrosol, 5-HMF, cytidine and uridine. The results showed no significant variations on the normalized areas of each studied compound, as a function of days of storage and storage conditions (during 90, stored at both -20 and -70 ºC) with percentage variations showing values < 6%, thus indicating that no significant samples degradation occurred during storage.

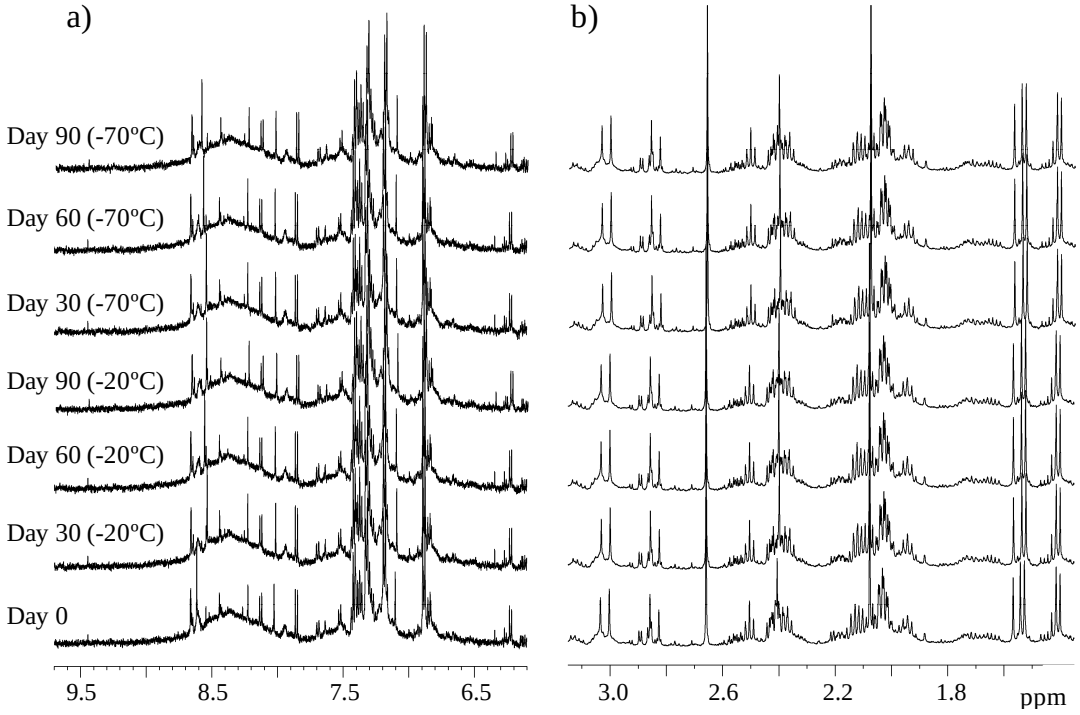

Figure 3.2. Examples of 500 MHz ^{1}H NMR spectra of a beer sample with different storage conditions (at -20ºC and -70 ºC, during 90 days), with insets shown for a) aliphatic (3.1-1.4 ppm) and b) aromatic (9.7-6.0 ppm) regions.

3.4 Beer characterization by 1D and 2D NMR and LC-NMR and LC-MS

Figure 3.3 shows a typical ^{1}H NMR spectrum of a lager beer, acquired at 500 MHz, with suppression of the water signal (4.8 ppm), where three distinct spectral regions containing significant peak density are identified: aliphatic (0-3 ppm), sugar (3-6 ppm) and aromatic (6-10 ppm) regions. For spectral assignment, 2D NMR spectra (i.e. TOCSY, HSQC and *J-resolved* spectra) were recorded, with examples shown in Figures 3.3, 3.4 and 3.5, respectively. These 2D spectra allowed a more detailed characterization, thus enabling the identification of 40 compounds, including amino and organic acids, higher alcohols and nucleosides, present in beer (Table 3.2). For each

studied brewing site (sites **A**, **B** and **C**), a beer sample was chosen for detailed NMR characterization, with the different beers showing a similar overall spectral feature.

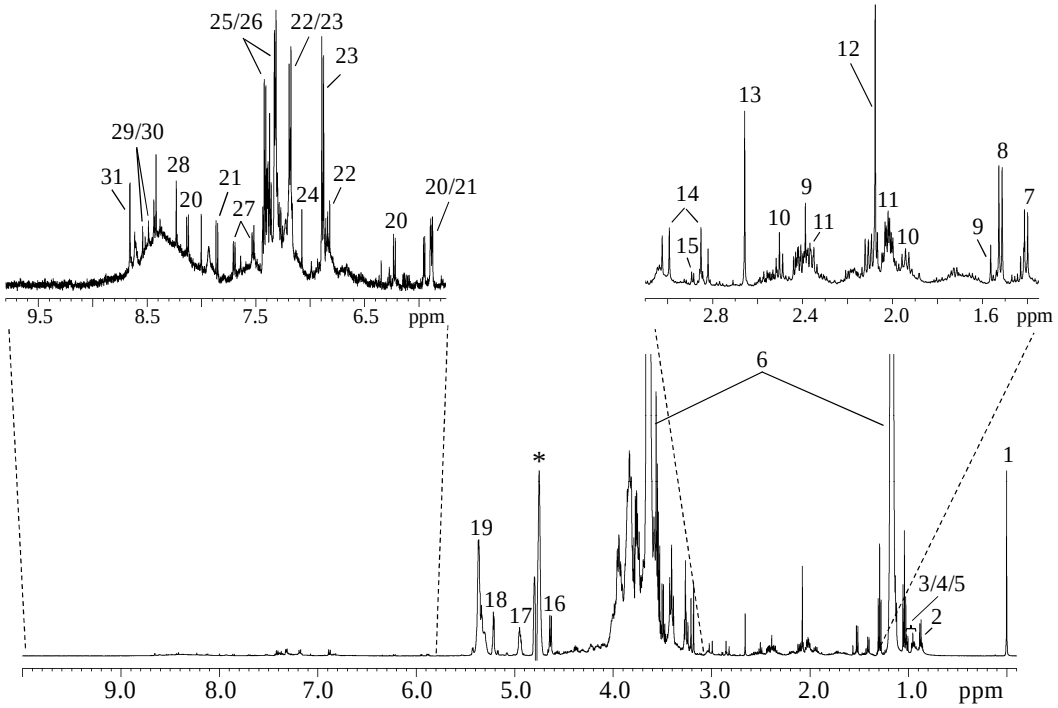

Figure 3.3. 500 MHz ^{1}H NMR spectrum of beer, with insets shown for aliphatic (0-3 ppm) and aromatic (6-10 ppm) regions. Peak assignments: 1, TSP (internal standard); 2, higher alcohols (isobutanol, isopentanol, propanol); 3, isoleucine; 4, leucine; 5, valine; 6, ethanol; 7, lactic acid; 8, alanine; 9, pyruvic acid and hydrate form; 10, γ-aminobutyric acid (GABA); 11, proline; 12, acetic acid; 13, succinic acid; 14, malic acid; 15, citric acid; 16-19, anomeric protons (H1): β red., α (1→6), α red., α (1→4), respectively; 20, cytidine; 21, uridine; 22, tyrosol; 23, tyrosine; 24, gallic acid; 25, phenylalanine; 26, phenylethanol; 27, tryptophan; 28, formic acid; 29, adenosine; 30, inosine; 31, histidine. * suppressed water resonance.

The aliphatic region (0-3 ppm) shows many signals of weak intensity and significantly overlapped (Figures 3.3 and 3.4b), comprising signals of ethanol, amino acids (alanine, arginine, GABA, glutamine, isoleucine, leucine, lysine, proline and valine), higher alcohols (isobutanol, isopentanol and propanol) and organic acids (acetic, citric, lactic, malic, pyruvic and succinic acids).

The sugar region (3-6 ppm) comprises signals from fermentable sugars (e.g. maltose and trehalose) and dextrins (Figure 3.3). When compared with the other spectral sub-regions, the sugar region shows lower resolution, reflecting the predominance of high-molecular weight oligomers, which undergo slower molecular tumbling leading to

faster transverse relaxation and, thus broader signals. Spectral distinction between the different sized oligomers present in beer is difficult, however, partial assignment of their signals may be carried out in the anomeric region (4.5-5.5 ppm), based on previous reports (Jodelet *et al.*, 1998; Vinogradov and Bock, 1998; Duarte *et al.*, 2003). The H1 protons corresponding to the reducing glucose ring of α- and β-anomers give rise to resonances at 5.20 and 4.64 ppm, respectively (Figure 3.3, peaks 18 and 16). The non-reducing anomeric H1 protons engaged in branched α (1→6) and in linear α (1→4) glycosidic linkages resonate at about 4.96 and 5.36 ppm, respectively (Figure 3.3, peaks 17 and 19). The latter resonance imparts the group of overlapped signals, in the 5.3-5.4 ppm range, revealing the presence of different-sized linear glucose segments.

The aromatic region (6-10 ppm) is the one of lowest peak intensity. Contributions arising from aromatic amino acids (histidine, tryptophan, phenylalanine and tyrosine) and corresponding alcohols (phenylethanol and tyrosol), nucleosides (adenosine and/or inosine, cytidine and uridine), gallic and formic acids and other characteristic compounds such as acetaldehyde and 5-HMF are identified, the latter two compounds being important indicators of beer staling (Guido *et al.*, 2003; Vanderhaegen *et al.*, 2003; Vanderhaegen *et al.*, 2006). Broad humps are detected in the aromatic region (Figure 3.3), especially between 7.9 and 8.7 ppm, resulting from aromatic compounds with large molecular weight (polyphenols) (Duarte *et al.*, 2004).

Furthermore, a high number of resonances are still unassigned, with emphasis in three intense pairs of ^{1}H-^{1}H correlations detected in the TOCSY spectrum: 6.51 and 7.46 ppm (weak intensities signals); 7.30 (overlapped with phenylalanine and phenylethanol signals in ^{1}H NMR) and 8.61 (broad hump) ppm, represented in Figure 3.4a by S1 and S2, respectively; and 2.89 (overlapped with malic acid signal in ^{1}H NMR) and 3.83 ppm (Figure 3.4b, by S3).

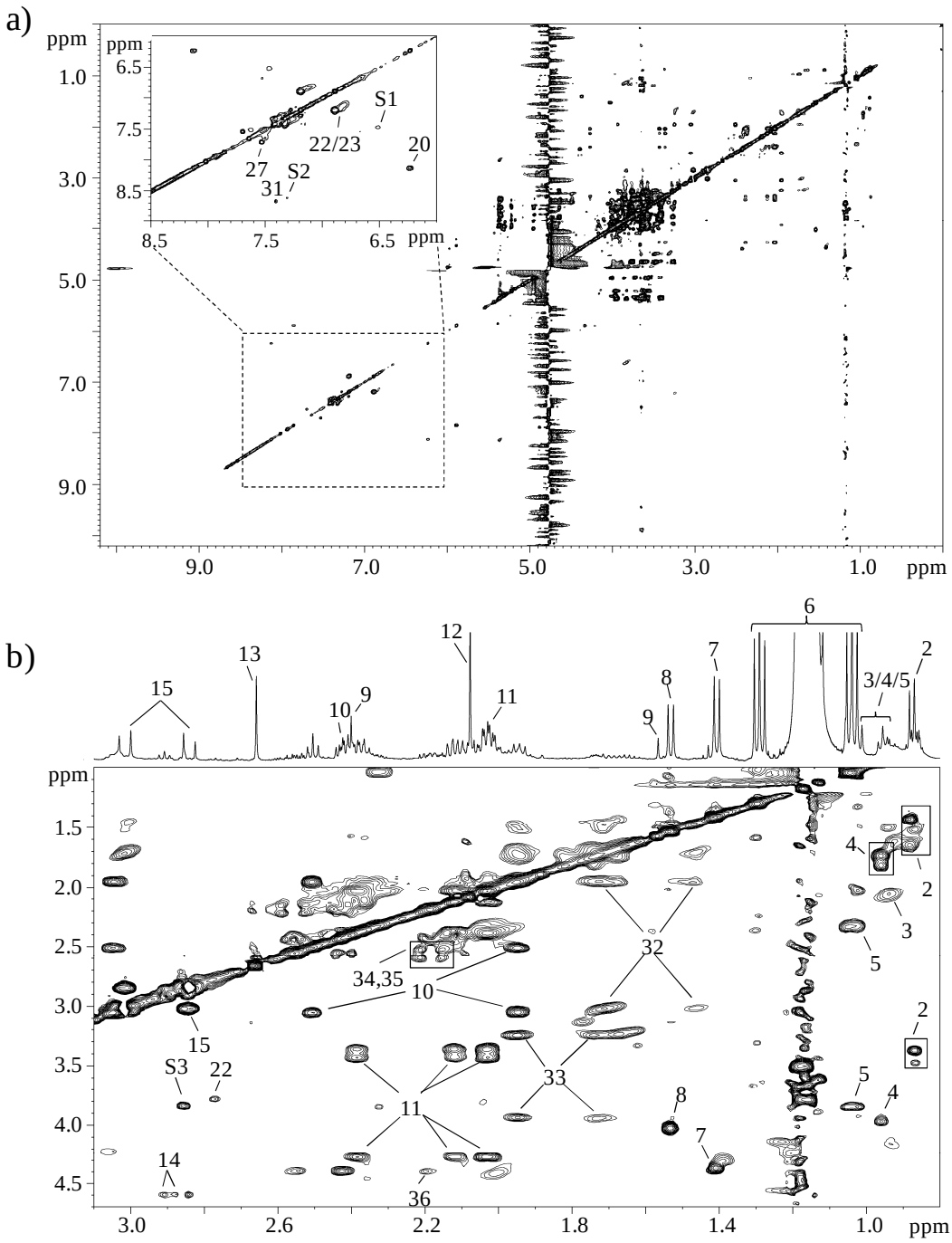

Figure 3.4. a) 500 MHz TOCSY spectrum of a lager beer, with inset shown for aromatic region (6.0-9.0 ppm); and b) aliphatic region (1.0-3.1 ppm). Peak assignments correspond to that shown in Figure 3.2, with the addition of 32, lysine; 33, arginine; 34, glutamic acid; 35, glutamine; 36, possibly acetoin. S1, S2 and S3 refer to unassigned pairs of resonances: 6.51 and 7.46 ppm; 7.30 and 8.61 ppm; and 2.89 and 3.83 ppm, respectively.

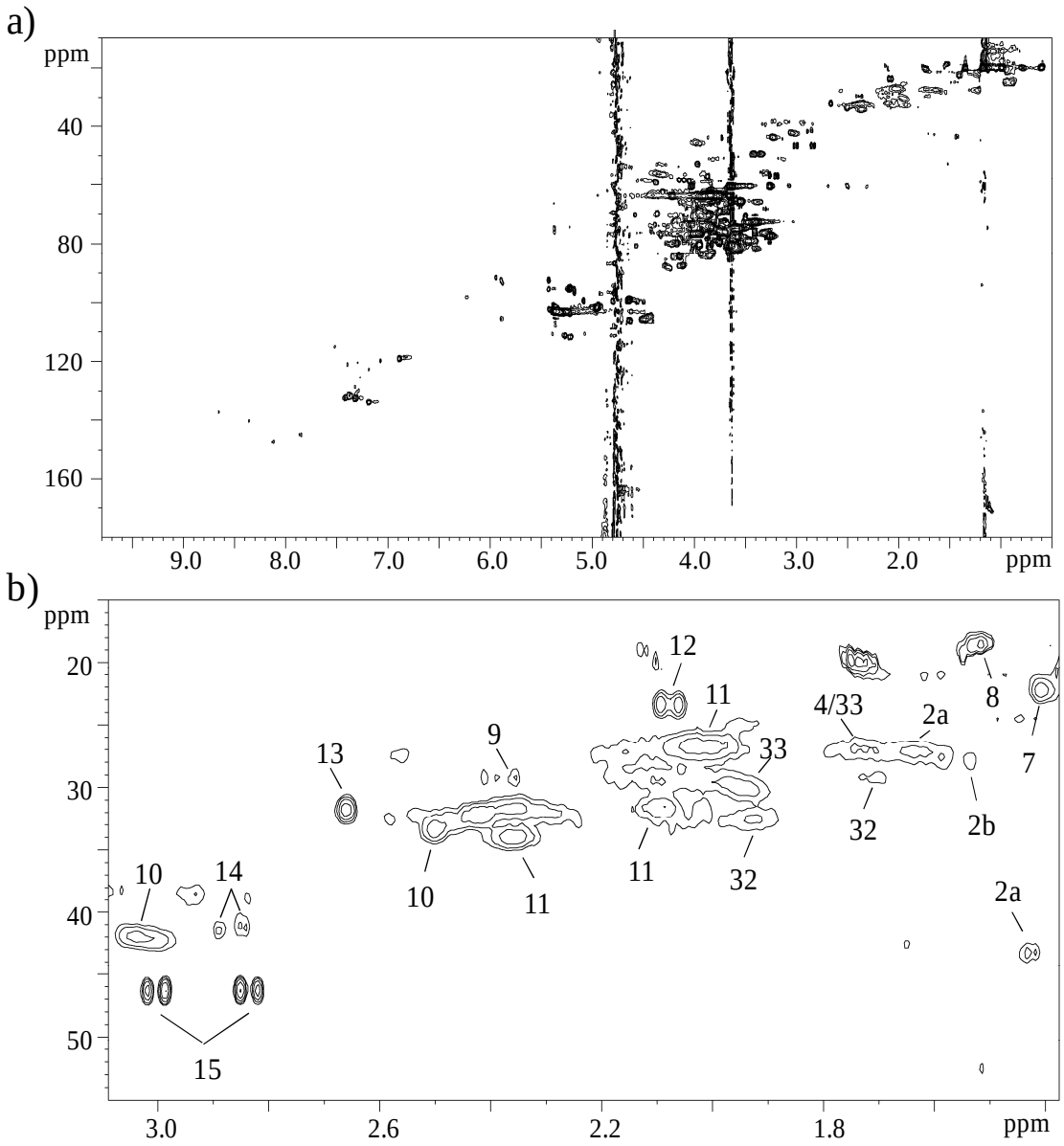

Figure 3.5. a) 500 MHz (125.8 MHz for ^{13}C) HSQC spectrum of a lager beer, with inset shown for b) aliphatic region (1.0-3.1 ppm). Peak assignments correspond to that shown in Figure 3.2, with the addition of 2a and 2b corresponding to isopentanol and propanol (higher alcohols).

The higher alcohols, isobutanol, isopentanol and propanol (especially the latter two), show significant overlapping in the ^{1}H NMR, but the analysis of the HSQC spectrum allowed the signals from isopentanol and propanol to be distinguished (Figure 3.5, peaks 2a and 2b).

The *J-resolved* spectrum enabled the detection of a slight separation between amino acid tyrosine and its corresponding alcohol, tyrosol (compounds almost

completely overlapped in 1D NMR, TOCSY and HSQC spectra), in the peaks arising at 6.88 and 6.85 ppm respectively (Figure 3.6b, identified as: 22 for tyrosol and 23 for tyrosine). Furthermore, multiplicities were confirmed for several compounds, e.g. cytidine (showing two doublets at 6.28 and 8.13 ppm (J = 7.9 Hz)); gallic acid (singlet at 7.09 ppm); and uridine (show three doublets, at 5.88 (J = 4.0 Hz), 5.89 (J = 8.1 Hz) and 7.86 (J = 8.1 Hz) ppm).

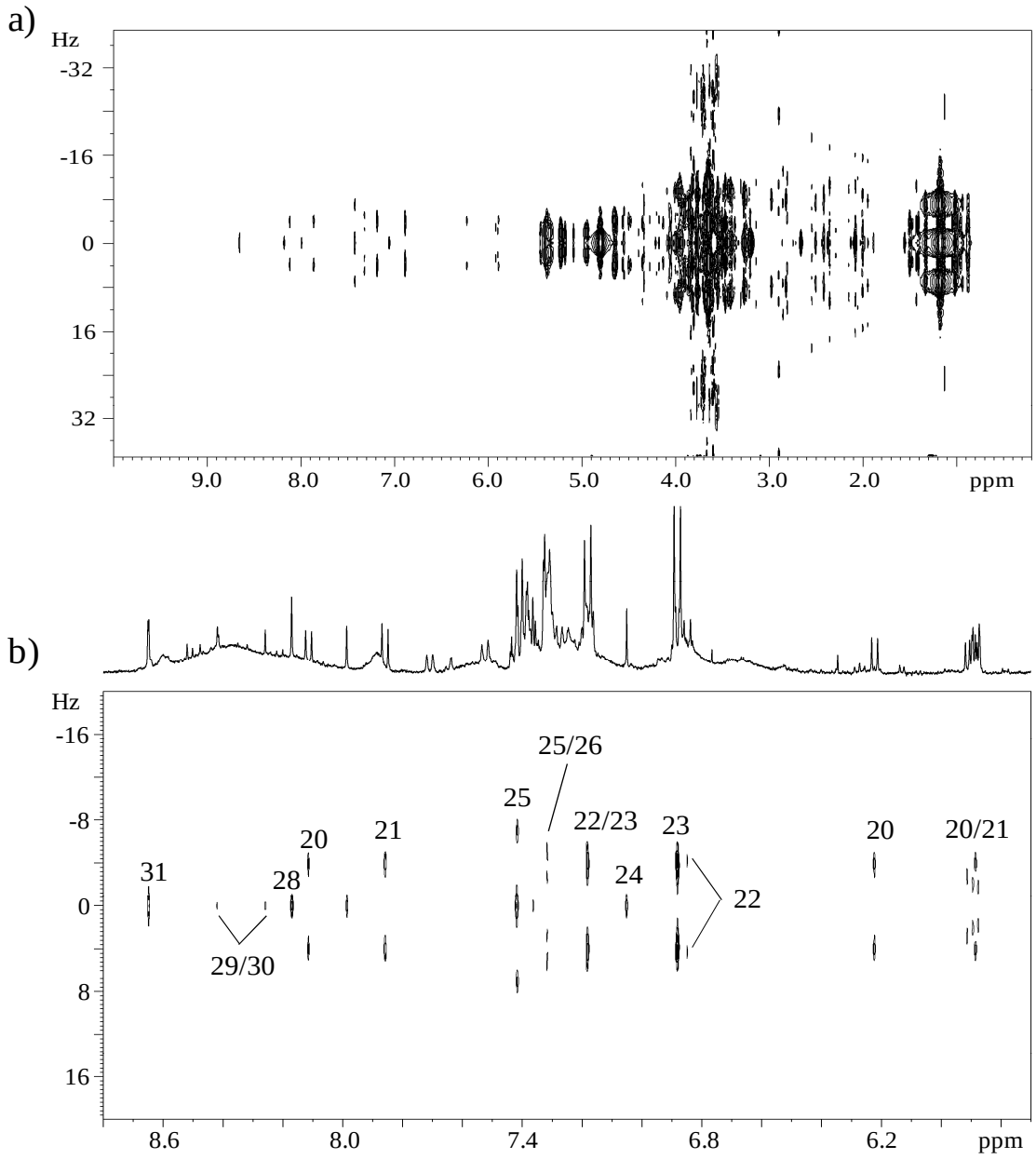

Figure 3.6. a) 500 MHZ 2D *J-resolved* spectrum of a lager beer, with inset shown for b) aromatic region (6.0-9.0 ppm). Peak assignments correspond to that shown in Figure 3.2.

In this way, a total of 40 compounds were identified in beer by analysis of 1D and 2D NMR spectra, with the complete list of assignments, ^{1}H and ^{13}C chemical shifts, multiplicity and J_{HH} coupling values shown in table 3.2. Although no new compounds were assigned when compared with literature, detailed information about beer when working at pH 1.9 (not encountered in both literature and data bases) was obtained. Examples of some structures of identified compounds present in beer are shown in Figure 3.7. However, there were still more than 20 spin systems and/or uncorrelated/ single resonances in the NMR spectra that remain unknown, thus requiring further investigation.

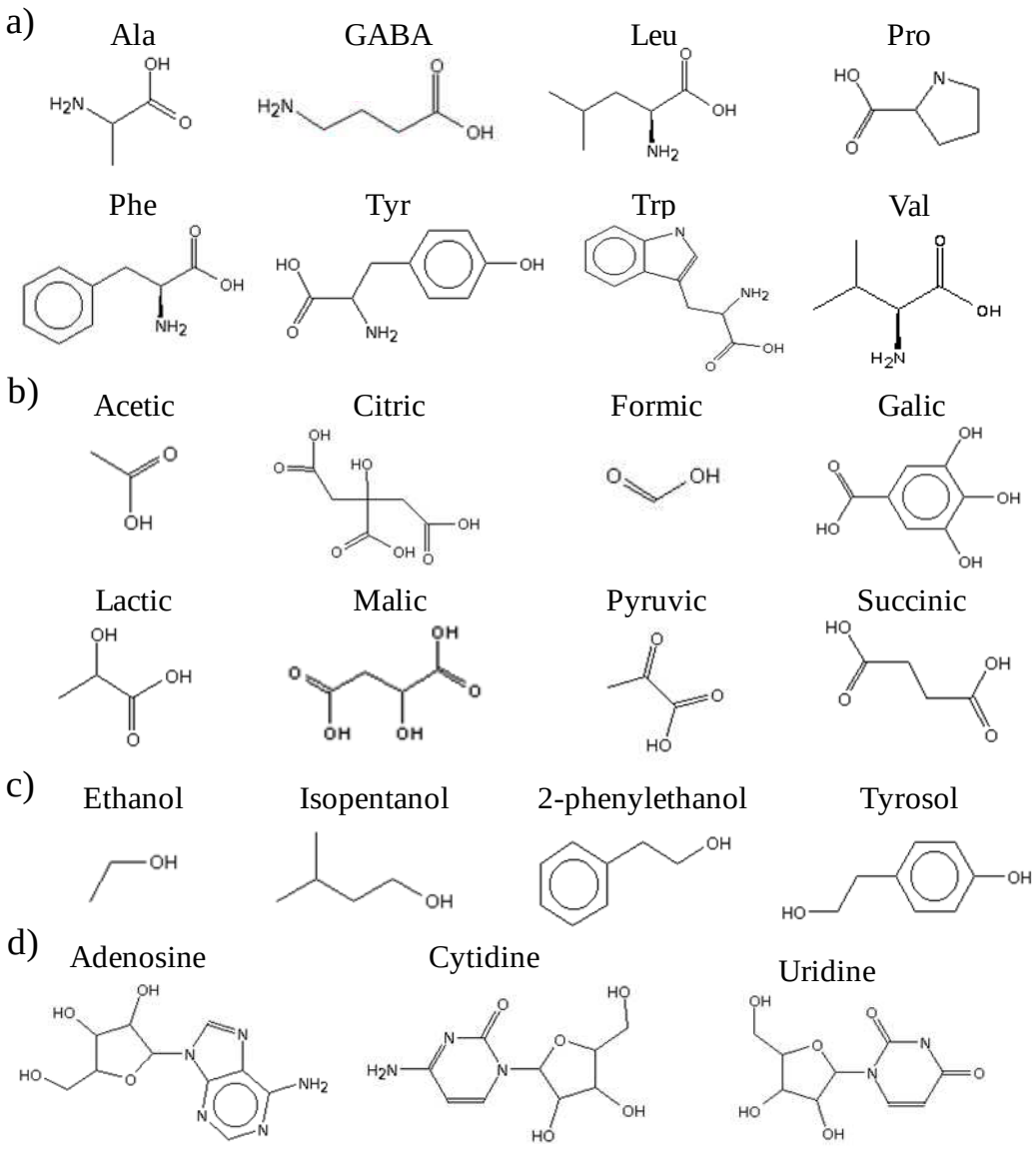

Figure 3.7. Structures of some compounds identified in beer by 1D and 2D NMR spectroscopy: a) amino acids (represented with the three letter code for amino acids), b) organic acids, c) alcohols and d) nucleosides.

Table 3.2. Compounds identified in beer and respective assignments, ^{1}H and ^{13}C chemical shifts, multiplicity of the proton signal and J_{HH} values. s: singlet, d: doublet, t: triplet, q: quartet, dd: doublet of doublets, m: multiplet; [a] tentative assignment; [b]: compounds for which a standard solution ^{1}H NMR spectrum was acquired (at pH 1.90).

Compounds	Assignments	^{1}H (δ) / ppm	*J* coupling/ Hz	^{13}C (δ) / ppm
Anomeric protons	H1α red.	5.20, d	3.6	94.9
	H1β red.	4.65, d	7.9	98.8
	H1 α (1→ 4)	5.37		102.6
	α (1→ 6)	4.95		101.4
Acetaldehyde [b]	CH_3	2.23, d	3.0	
	CH	9.66, q	3.0	
Acetic acid [b]	β CH_3	2.08, s		23.3
Acetoin [a]	CH_3	1.36, d		
	CH_3	2.21, s		
	CH	4.41, q		
Adenosine and/or	CH_2	3.84, dd	4.4; 8.4	
Inosine [b]	CH_2	3.92, dd	3.0; 9.7	
	C4'H, ribose	4.26, q	3.9	
	C3'H, ribose	4.43, t	4.9	
	C1'H, ribose	6.14, d	5.3	
	C8H, ring	8.44, s		
	C2H, ring	8.54, s		
Alanine	β CH_3	1.52, d	7.3	18.4
	α CH	3.88, q		52.6
γ-aminobutyric acid	β CH_2	1.94, m		
(GABA) [b]	α CH2	2.50. t	7.2	33.4
	γ CH_2	3.04, t		42.1
Arginine [b]	γ CH_2	1.66, m		27.0
	β CH2	1.92, m		29.8
	δ CH_2	3.24, t		43.4
	α CH	3.83, t	6.4	
Citric acid [b]	α, γ CH_2	2.83, d	15.8	46.3
	α', γ' CH_2	3.0, d	15.8	
Cytidine	CH_2	3.80, dd		
	CH_2	3.93, m		
	C4H, ribose	4.16, m		
	C3H, ribose	4.19, t		
	C2H, ribose	4.33, t		
	C1H, ribose	5.87, d		92.3
	C6H, ring	6.22, d	7.9	98.0
	C5H, ring	8.13, d	7.9	147.1
Ethanol	CH_3	1.17, t	7.1	19.5
	$\underline{CH_2}OH$	3.64, q	7.1	60.9
Ethyl acetate [a,b]	CH_3	1.24, t	7.1	
	CH_3	2.06, s		
	$O\underline{CH_2}$	4.13, q	7.1	
Formic acid [b]	$\underline{H}COOH$	8.23, s		
Gallic acid	C2H, C6H ring	7.07, s		119.7
Glutamic acid [b]	β, β' CH_2	2.15, m		
	γ CH_2	2.56, m		32.4
	α CH	3.86, t	6.4	

Table 3.2 (cont.)

Compounds	Assignments	^{1}H (δ) / ppm	*J* coupling/ Hz	^{13}C (δ) / ppm
Glutamine [b]	β CH_2	2.18, m		
	γ CH_2	2.49, m		33.4
	α CH	3.86, t	6.3	
Glycerol [a]	CH_2	3.55, dd	6.4; 5.3	
	CH_2	3.64, dd		65.4
	CH	3.77, m		
Histidine [b]	β, β' CH	3.37, dd		
	α CH	4.12, dd		
	C4H, ring	7.39, s		120.8
	C2H, ring	8.66, d		136.9
5-HMF	CH_2	4.70, s		
	C2H, ring	6.67, d	3.9	
	C3H, ring	7.52, d	3.9	
	CHO	9.45, s		
Isobutanol [b]	CH_3	0.86, d	6.6	13.2
	CH	1.73, m		
	$\underline{CH_2}OH$	3.36, d	6.6	
Isoleucine [b]	δ CH_3	0.95, t		
	γ' CH_3	1.01, d	7.0	
	γ CH	1.32, m		27.4
	γ' CH	1.47, m		
	β CH	2.10, m		
	α CH	3.93,m		
Isopentanol [b]	CH_3	0.88, d	6.6	24.8
	CH_2	1.42, q	6.9	43.3
	CH	1.64, m		26.8
	$\underline{CH_2}OH$	3.63, t	6.9	
Lactic acid [b]	β CH_3	1.40, d	7.0	22.3
	α CH	4.37, q	7.0	69.5
Leucine [b]	δ, δ' CH_3	0.96, t	6.3	24.6
	γ CH	1.70, m		26.8
	β CH_2	1.80, m		42.2
	α CH	3.95, t		
Lysine	γ CH_2	1.52, m		24.6
	δ CH_2	1.73, m		29.3
	β CH_2	1.92, m		32.7
	ε CH_2	3.02, t		41.9
	α CH	3.79, t		
Malic acid [b]	β' CH	2,89, dd	6.8; 9.7	41.1
	β CH	2,92, dd	4.5; 11.9	41.4
	α CH	4,58, dd	4.7	69.4
Maltose		5.43		92.4
Phenylalanine [b]	β' CH	3.17, dd	7.8; 6.7	
	β CH	3.32, dd	5.4; 9.1	39.0
	α CH	4.16, dd	5.4; 2.4	58.5
	C2H, C6H, ring	7,32, d	7.5	131.9
	C4H, ring	7,38, m		131.1
	C3H, C5H, ring	7,42, m		131.9
2-Phenylethanol [b]	CH_2	2.85, t	6.7	
	CH_2	3.83, t	6.7	
	C2H, C6H, ring	7.31, m		129.4
	C3H, C5H, ring	7.38, m		132.0

Table 3.2 (cont.)

Compounds	Assignments	^{1}H (δ) / ppm	*J* coupling/ Hz	^{13}C (δ) / ppm
Proline [b]	γ CH_2	2.02, m		26.7
	β' CH	2.11, m		31.6
	β CH	2.37, m		31.7
	δ' CH	3.35, t		49.1
	δ CH	3.41, t		
	α CH	4.23, t		
Propanol [b]	CH_3	0.88, t	7.5	13.3
	CH_2	1.53, m		27.9
	$\underline{CH_2}OH$	3.54, t	6.7	
Pyruvic acid and		1.56, s		
hydrate form [b]	β CH_3	2.40, s		29.3
Succinic acid [b]	α, β CH_2	2,65, s		31.5
Threonine	γ CH_3	1.29, d		
	α CH	3.65, d		
	β CH	4.25, m		
Tryptophan	β CH	3.34, dd		
	β' CH	3.49, dd		
	α CH	4.19, dd		
	C5H, ring	7.18, m		122.5
	C6H, ring	7.27, m		125.2
	C2H, ring	7.31, s		128.3
	C7H, ring	7.53, d	8.0	114.8
	C4H, ring	7.70, d	8.0	121.2
Tyrosine [b]	β' CH	3.10, dd	7.8; 7.0	38.0
	β CH	3.23, dd	5.4; 9.2	
	α CH	4.12, dd	5.4; 2.2	
	C3H, C5H, ring	6.88, d	8.5	118.9
	C2H, C6H, ring	7.18, d	8.5	133.6
Tyrosol [b]	β CH_2	2.77, t	6.7	39.4
	α CH_2	3.77, t	6.7	
	C3H, C5H, ring	6.85, d	8.5	118.6
	C2H, C6H, ring	7.17, d	8.5	133.6
Trehalose	C1H, ring	5.18, d		96.0
Uridine	CH	3.79, dd		
	CH	3.89, dd		
	C4'H, ribose	4.11, dd		
	C3'H, ribose	4.21, t		
	C2'H, ribose	4.33, t		
	C5H, ring	5.88, d		105.3
	C1'H, ribose	5.89, d		92.1
	C6H, ring	7.86, d	8.2	144.7
Valine	γ CH_3	0,96, d		20.0
	γ' CH_3	1.01, d		20.5
	β CH	2.26, m		
	α CH	3.64, d		

In order to further characterize beer composition, LC-NMR combined with LC-MS was used. Figure 3.8 shows the UV chromatogram obtained for a lager beer sample, using a reverse phase column for the detection of less-polar analytes. In this study, the

same conditions as the ones used in the characterization of aromatic compounds in some liquid foods, including juices, beer and wine (Gil *et al*., 2003) were applied to beer characterization. As seen in Figure 3.8, a large number of peaks can be visually observed, although many of which are overlapped, reflecting the existence of significant co-elution of the different analytes in the beer matrix. In cases where poor chromatographic separation occurs, the relevant compounds under investigation have weak UV chromophores or if the exact chromatographic retention time is poorly defined, the LC-NMR method of time-slice is commonly used. By combining the acquired 1D ^{1}H NMR spectra as a function of chromatographic time (each spectrum set as a row), a pseudo on-flow diagram may be obtained (Figure 3.9), facilitating the visualization/ interpretation of the obtained data. LC-MS experiments were performed separately, in a continuous flow chromatographic run, although using the same chromatographic conditions. Overall, LC-NMR, combined with LC-MS, data enabled the identification of 20 compounds in beer (Table 3.3), however no new compounds, compared with those identified by 1D and 2D NMR, was achieved.

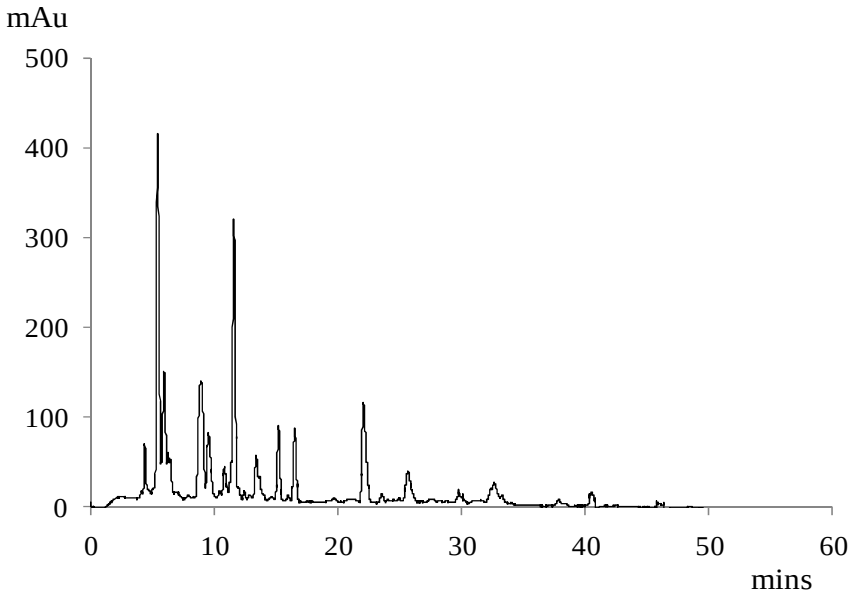

Figure 3.8. Lager beer chromatogram acquired at 280 nm, using a reverse phase column.

Figure 3.9 illustrates the pseudo on-flow NMR chromatogram based on the time-sliced method. At early retention times (R.T.), high abundance of bands is observed, reflecting the compounds co-elution mentioned before. In fact, the majority of the identified compounds (16 in a total of 20) elute before 30 minutes, comprising the most

polar metabolites, namely amino acids (e.g. arginine, GABA, isoleucine, histidine, leucine, phenylalanine, tryptophan, tyrosine and valine) nucleosides (cytidine and uridine), organic acids (citric, lactic, malic, pyruvic and succinic acids), ethanol and sugars. The first observed compounds are identified in row 12, corresponding to R.T 6.0 minutes (Figure 3.9), and showing the characteristic spin systems of arginine, GABA, and histidine. Observing the positive mode mass spectrum, obtained by LC-MS, and analyzing the same chromatographic region, one can identify peaks arising at *m/z* values 174 and 104 corresponding to the molecular ion, $[M+D]^+$, of arginine (Mr 173) and GABA (Mr 103), respectively.

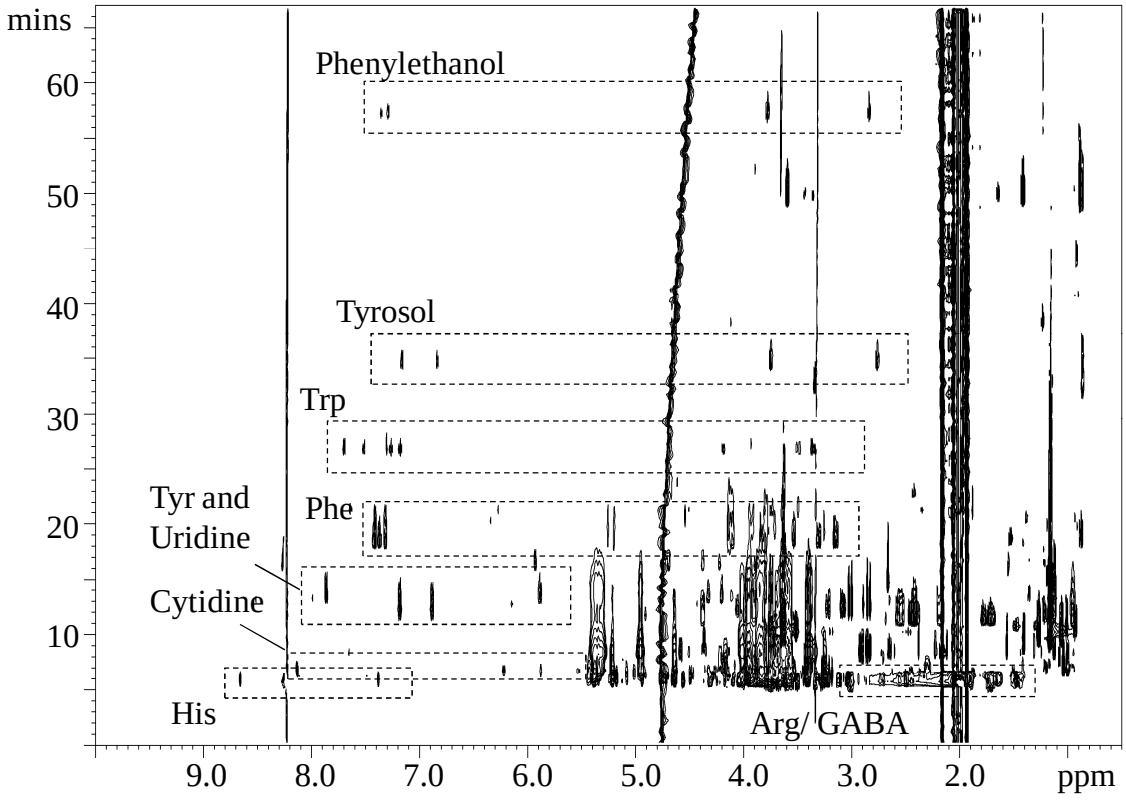

Figure 3.9. Pseudo on-flow NMR chromatogram acquired in time-slice mode, with some assignments. Dashed shapes were drawn to help visualization.

Following the same analysis procedure, the following compounds were also assigned: cytidine and valine (both co-eluting at R.T. 6.9, and $[M+D]^+$ of 244 and 118, respectively), lactic, malic and pyruvic acids (co-eluting at R.T. 8.6 minutes,), ethanol (R.T. 10.0 minutes), isoleucine and leucine (co-eluting at R.T. 11.1 minutes and showing $[M+D]^+$ 132), citric and tyrosine (co-eluting at R.T. 12.1 minutes), uridine (R.T. 13.3 minutes), succinic acid (R.T. 14.5 minutes), phenylalanine (R.T. 18.2

minutes and $[M+D]^+$ 166) and tryptophan (R.T. 27.0 minutes and $[M+D]^+$ 205). Examples of ^{1}H NMR and MS spectral information used for metabolite identification are shown in Figure 3.10 for a) cytidine and valine, b) phenylalanine and c) tryptophan. Furthermore, the carbohydrates also elute at the beginning of the chromatography (until 25 minutes), as seen in Figure 3.9, originating a high complexity and number of signals arising in the region between 3.0-5.5 ppm. After 30 minutes, the pseudo on-flow NMR chromatogram shows much "cleaner", with only three bands being identified on the pseudo on-flow, reflecting the spin systems of three alcohols, i.e. tyrosol (R.T. 34.3 minutes), isopentanol (R.T. 49.7 minutes) and phenylethanol (57.2 minutes).

Table 3.3. MS and NMR data obtained for a lager beer composition obtained by LC-NMR combined with LC-MS. n.i. refers to not identified MS fragment for that metabolite.

R.T./ mins	MS (*m/z*)	NMR signals ^{1}H (δ) / ppm (multiplicity)	Compound
6.0	174	1.66 (m), 1.72 (m), 1.92 (m), 3.23 (t), 3.83 (t)	Arginine
	104	1.94 (m), 3.50 (t), 3.04 (t)	GABA
	n.i.	7.39 (s), 8.66 (d)	Histidine
6.9	244	4.33 (t), 5.87 (d), 6.22 (d), 8.13 (d)	Cytidine
	118	0.96 (d), 1.01 (d), 2.26 (m)	Valine
8.6	n.i.	1.40 (d), 4.36 (q)	Lactic acid
	118	2.85 (dd), 2.89 (dd)	Malic acid
	n.i.	1.56 (s), 2.39 (s)	Pyruvic acid
10.0	n.i.	1.17 (t), 3.64 (q)	Ethanol
11.1	132	0.95 (t), 1.01 (d), 1.47 (m)	Isoleucine
	132	0.96 (t), 1.70 (m), 1.80 (m), 3.95 (t)	Leucine
12.1	n.i.	2.83 (d), 3.00 (d)	Citric
	n.i.	3.10 (dd), 3.23 (dd), 4.12 (q), 6.88 (d), 7.18 (d)	Tyrosine
13.5	n.i.	4.21 (t), 4.33 (t), 5.88 (d), 7.86 (d)	Uridine
14.5	n.i.	2.65 (s)	Succinic acid
18.2	166	3.17 (dd), 3.32 (dd), 4.16 (dd), 7.38 (m), 7.42 (m)	Phenylalanine
27.0	205	3.34 (dd), 3.49 (dd), 4.19 (dd), 7.18 (m), 7.27(m), 7.31 (s), 7.53 (d), 7.70 (d)	Tryptophan
34.3	n.i.	2.77 (t), 3.77 (t), 6.85 (d), 7.17 (d)	Tyrosol
49.7	n.i.	0.86 (d), 1.4 (q), 1.63 (m), 3.61 (t)	Isopentanol
57.2	n.i	2.85 (t), 3.83 (t), 7.31 (m), 7.38 (m)	Phenylethanol

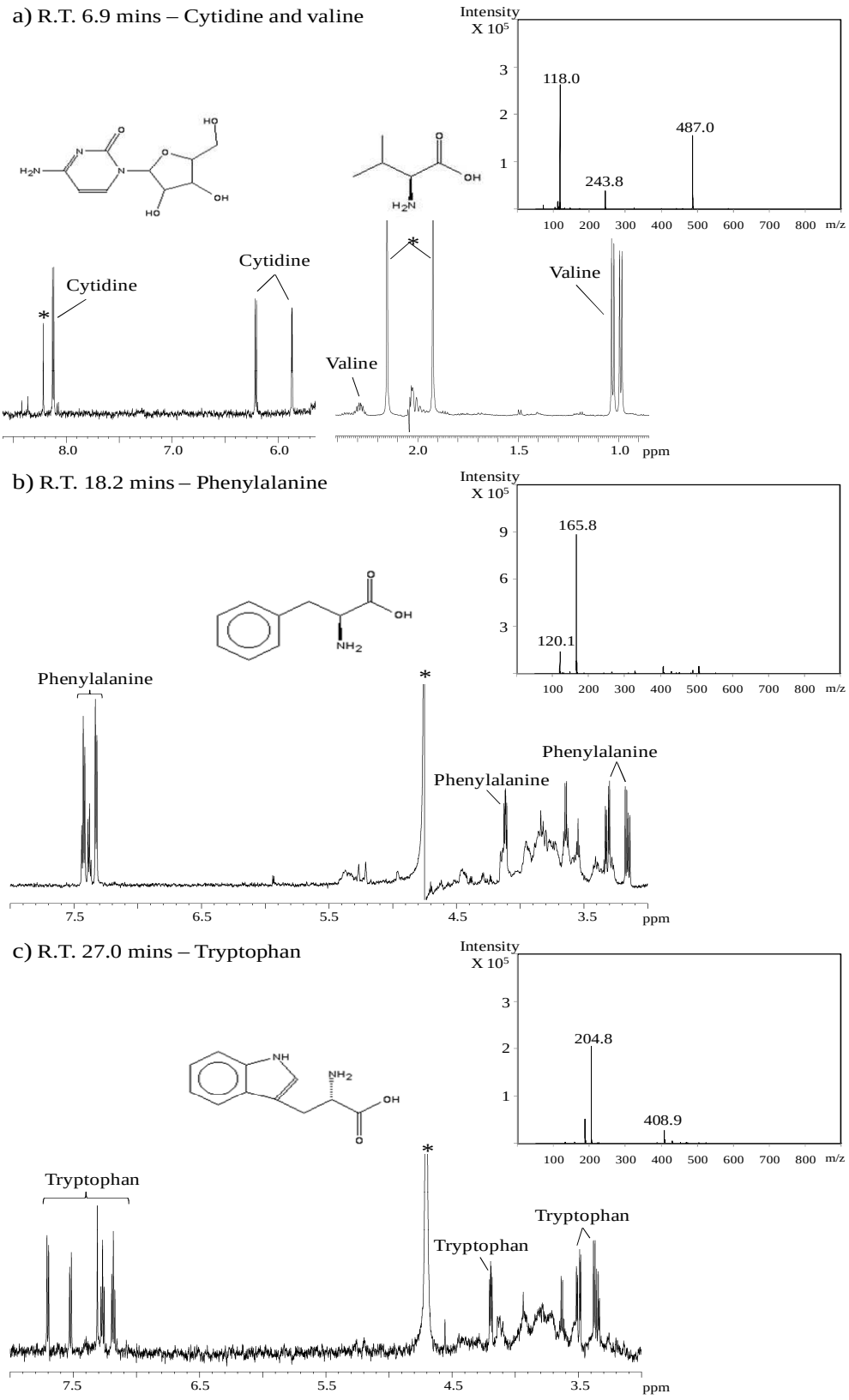

Figure 3.10. ^{1}H NMR and MS spectra corresponding to a) R.T. 6.9, cytidine and valine; b) R.T. 18.2 mins, phenylalanine; and c) R.T. 27.0, tryptophan. * refers to eluent peaks: acetonitrile (2.0 ppm), water (4.71 ppm) and formic acid (8.23 ppm).

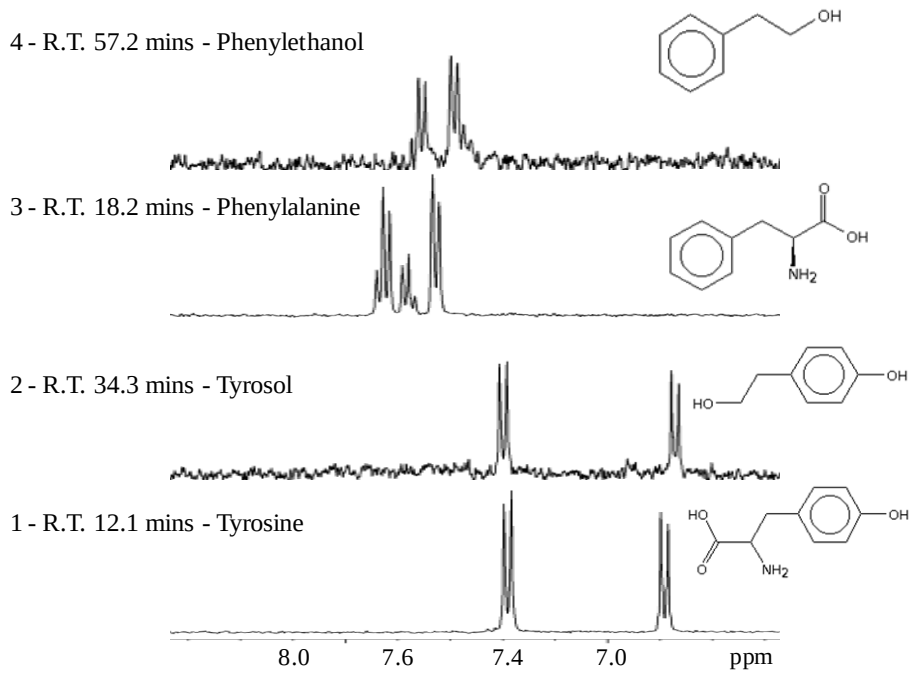

Figure 3.11. Rows of the pseudo on-flow NMR chromatogram corresponding to: 1, tyrosine; 2, tyrosol; 3, phenylalanine; and 4, phenylethanol.

The LC-NMR methodology, combined with LC-MS, is a useful tool in the chemical characterization of complex matrices, as in beer, allowing achieving a clearer and more complete identification and assignment of compounds, even in the case of different metabolites with similar spin systems. An example is seen for the pairs of aromatic amino acids and corresponding alcohols: phenylalanine/ phenylethanol and tyrosine/ tyrosol. In 1D and 2D NMR spectra, those metabolites show similar assignment in the aromatic region, with tyrosine and tyrosol showing two doublets (δ at 6.88 and 7.18 ppm for tyrosine and 6.85 and 7.17 ppm for tyrosol), whereas phenylalanine (peaks at 7.32 ppm, doublet; and at 7.38 ppm, multiplet) and phenylethanol (peaks at 7.31 and 7.38 ppm, multiplets) overlap at the region at 7.27-7.40 ppm. Although a separation between tyrosine and tyrosol has been was already observed in the *J-resolved* experiments (Figure 3.6), by LC-NMR these same compounds are chromatographic separated (Figures 3.9 and 3.11), with the amino acids eluting earlier (tyrosine and phenylalanine at R.T. 12.1 and 18.2 minutes) than the corresponding alcohols (tyrosol and phenylethanol at R.T. 34.3 and 57.2 minutes), and therefore, easing the compounds assignment. These results are in agreement with a previous report, where the same compounds were also distinguished in beer by LC-NMR/MS (Gil *et al.*, 2003).

In the future, in order to obtain a more detailed chemical characterization of beer using LC-NMR, improvements in this methodology are required, such as achieving a better chromatographic separation between beer components, the use of the loop-storage method and the optimization of samples pre-concentration methods, which would lead to higher sensitivity of the method.

3.5 Multivariate analysis of NMR data for the evaluation of beer variability between national brewing sites and production dates

To investigate the main sources of variability between beers of the same brand produced in the same country but in different brewing sites and at production dates, PCA and PLS-DA models were developed considering initially the whole spectra (0.8-10.0 ppm). Figure 3.12 shows the two-dimensional (2D) PCA (in the left) and the corresponding PLS-DA (in the right) scores scatter plots obtained using mean centering data processing. Autoscaling of NMR data was also employed, however, no distinction between either brewing sites or production dates was obtained.

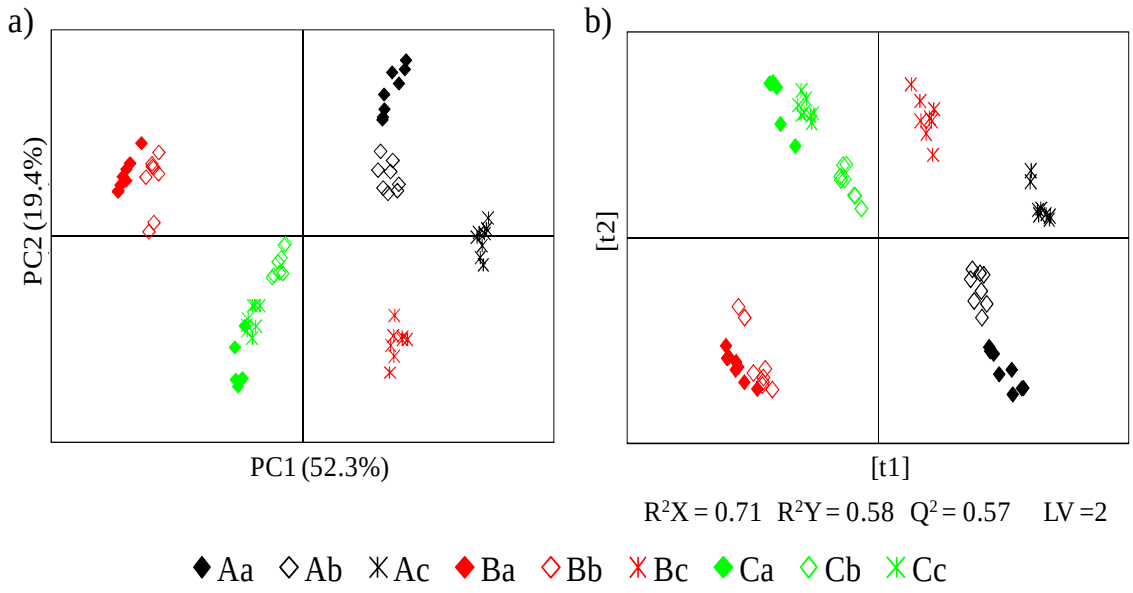

Figure 3.12. a) PCA and b) PLS-DA scores scatter plots obtained for whole NMR spectra as a function of brewing sites, described as: **A** (black), **B** (red) and **C** (green), and production dates, three different production dates described as: **a**, **b** and **c**.

The obtained PCA scores scatter plot (Figure 3.12a) comprises two main principal components (PC's), PC1 and PC2, which account for 52.3% and 19.4%, respectively, of the total variability. It shows that beers produced in site **A** (black) are positioned in positive PC1 and PC2, beers from site **C** (green) are in negative PC1 and PC2, but site **B** (red) shows some degree of variability between production dates with beers produced at date **c** (**Bc**), located at positive PC1 and negative PC2, being separated from the remaining beer samples produced in that site, positioned at negative PC1 and positive PC2. Figures 3.13a and 3.13b illustrate the PCA loadings profiles associated with PC1 and PC2, respectively. A major contribution from the carbohydrates region (between 3.0-5.5 ppm) is identified in both loading profiles, which may however mask potential changes in overlapped minor components. In Figure 3.13a, the positive signals reflect metabolites present in higher amounts in beer produced at sites **A** and **B**, date **c** (**Bc**). Accordingly, higher amounts in acetaldehyde, aliphatic and aromatic amino acids and linear dextrins (viewed by H1 resonances for β red., α red. and α (1→4) in linear oligomers) are suggested for site **A** and **Bc** beer. On the other hand, negative peaks correspond to metabolites present in higher quantities in beer from brewing sites **B** (except for **Bc**) and **C**, those signals being identified as higher and aromatic alcohols, nucleosides, organic acids, GABA, proline and branched dextrins (viewed by H1 resonances for α (1→6) and α (1→4) in branched oligomers). Furthermore, Figure 3.13b shows the PC2 loadings profile where sites **B** (except for **Bc**) and **C** can be distinguished, as beers from site **B** (positive PC2) are suggested having higher content of linear dextrins, aliphatic and aromatic amino acids, nucleosides and acetic, citric, formic, pyruvic and succinic acids, whereas beer from site **C** (negative side) are characterized by higher amounts of branched dextrins, proline and lactic and malic acids. These compositional trends are listed in Table 3.4.

For PLS-DA, the corresponding loadings profiles (not shown) were also investigated, showing the same spectral information that in PCA.

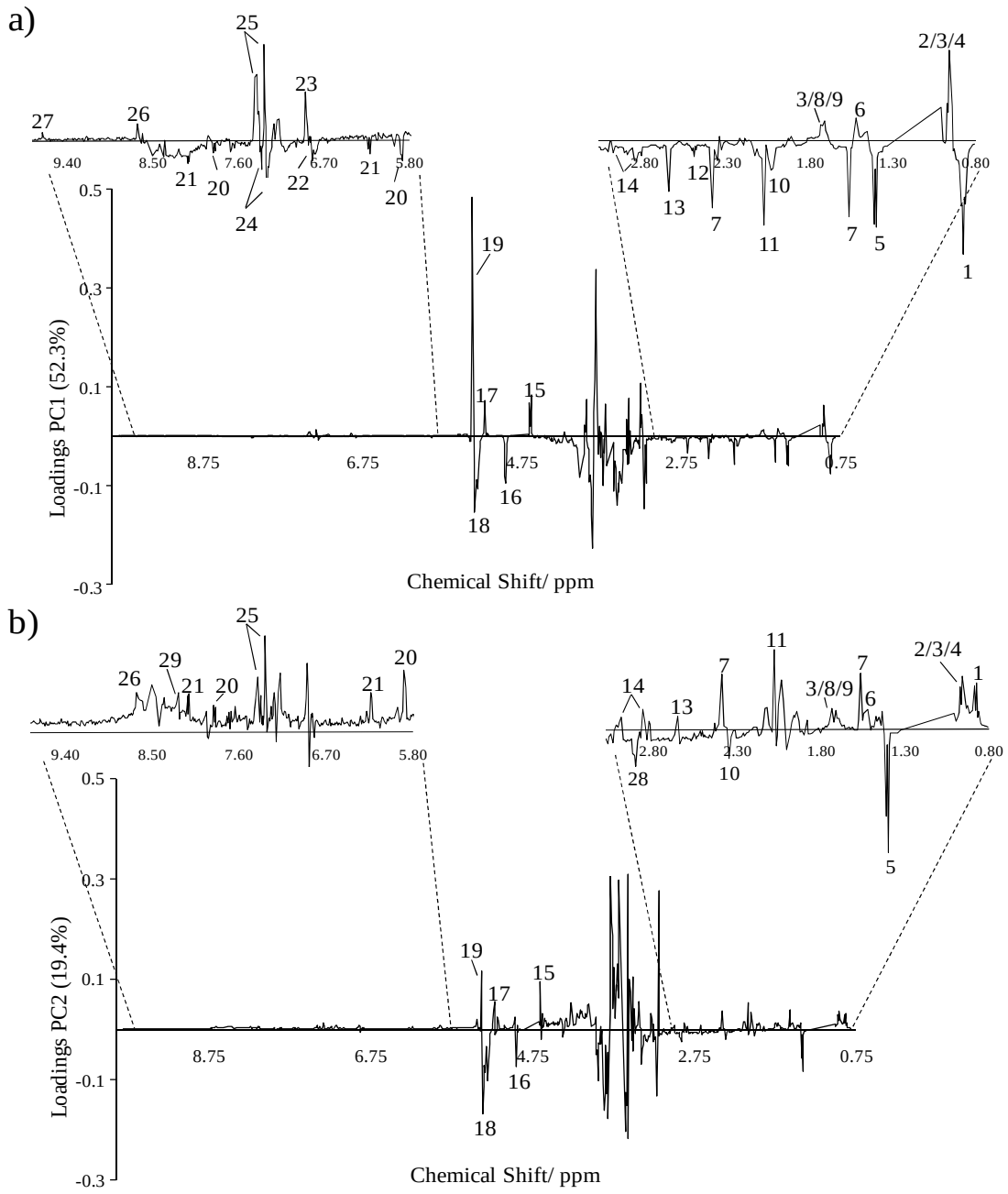

Figure 3.13. PCA loadings profiles corresponding to a) PC1 and b) PC2 obtained for whole NMR spectra. Signals identified refer to 1, higher alcohols (isobutanol, isopentanol and propanol); 2, isoleucine; 3, leucine; 4, valine; 5, lactic acid; 6, alanine; 7, pyruvic acid and hydrate form; 8, arginine; 9, lysine; 10, proline; 11, acetic acid; 12, GABA; 13, succinic acid; 14, citric acid; 15, H1β (red.); 16, H1α (1→6); 17, H1α (red.); 18, branched H1α (1→4); 19, linear H1α (1→4); 20, uridine; 21, cytidine; 22, tyrosol; 23, tyrosine; 24, phenylethanol; 25, phenylalanine; 26, histidine; 27, acetaldehyde; 28, malic acid; and 29, formic acid.

In order to confirm and improve this information, the aliphatic, anomeric and aromatic spectral regions were analyzed separately (results in Table 3.4). For the aliphatic region (0.8-3.1 ppm), the obtained scores scatter plot (Figure 3.14a), with PC1 and PC2 accounting for 42.9% and 24.7%, respectively, of the total variability, shows significant differences when compared with the one using the whole spectra (Figure 3.12a). Indeed, although beers from site **A** (black) show good reproducibility (negative PC1), beers produced in sites **B** (red) and **C** (green) show high variability within sites, thus reflecting lesser reproducibility of their brewing process in terms of aliphatic composition. On the basis of the results (Figure 3.14), it is suggested that beers from site **A**, and from **Bc** and **Cb** (negative PC1) are characterized by higher quantities of aliphatic amino acids, whereas beers from sites/dates **Bb**, **Cc** and **Ca** (positive PC1), the latter in less extent, have higher amounts of higher alcohols and organic acids (lactic and pyruvic). In addition, beers from site **Cc** (negative PC2) are separated from the ones produced in sites **Bc** and **Cb** (positive PC2), indicating that these beers are richer in aliphatic amino acids, lactic acid and an unassigned compound (triplet, at 2.90 ppm, described in Figure 3.14b and c with *), while beers from **Bc** and **Cb** beers show higher amounts of higher alcohols and organic acids (acetic, citric, pyruvic and succinic). In previous report related to brewing sites located at different countries (Almeida *et al*., 2006) only pyruvic and lactic acids were identified as varying within sites, mainly due to first derivative-like effects in the spectra. These effects were mitigated in this work by adjusting beer samples pH to acidic values (~1.90).

The majority of these compositional variations are in agreement with those observed in the whole spectra, with the exception of acetic acid, GABA and proline, for which no tendency was observed considering the aliphatic region alone. This apparently contradictory result possibly reflects the low magnitude of variation of those compounds. Furthermore, in beers from site **Cc**, an unassigned compound (triplet, at 2.90 ppm), overlapped with malic acid (double doublets, at 2.89 and 2.92 ppm), is noted.

Regarding the anomeric region (4.5-5.8 ppm), the PCA separation identified in the scores plot (Figure 3.15), with PC1 and PC2 accounting for 85.5% and 8.5%, respectively, of the total variability, is similar to the one seen for the whole spectra. This was expected since this region show significant higher intensity, thus dominating the separation observed in the whole spectra PCA model.

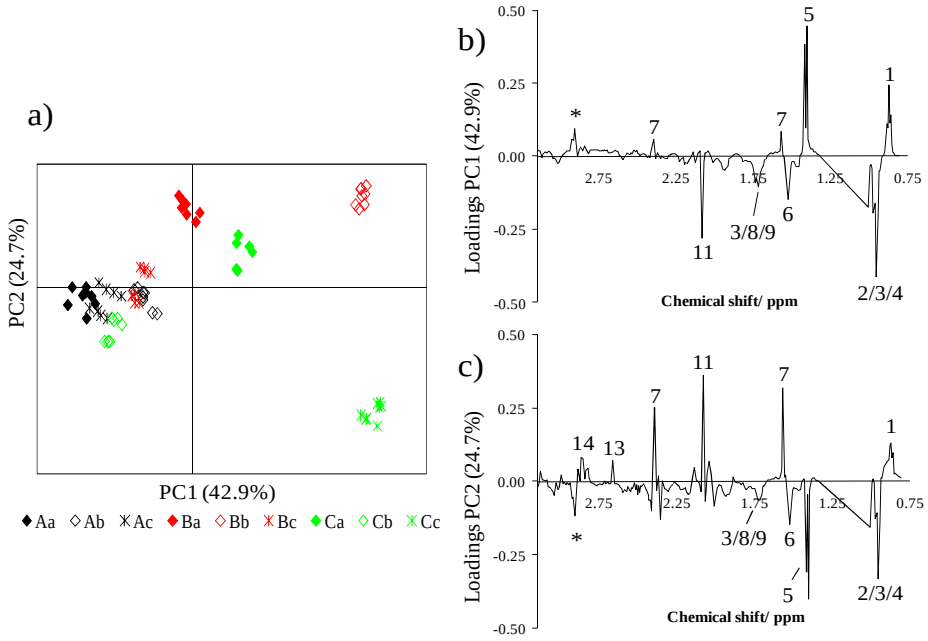

Figure 3.14. PCA for aliphatic region (0.8-3.1 ppm) as a function of brewing sites and production dates: a) PC1 vs. PC2 scores scatter plot, b) PC1 and c) PC2 loadings profiles. Signals varying are identified by the numbering shown in Figure 3.12. * unassigned signal (at 2.90 ppm, triplet).

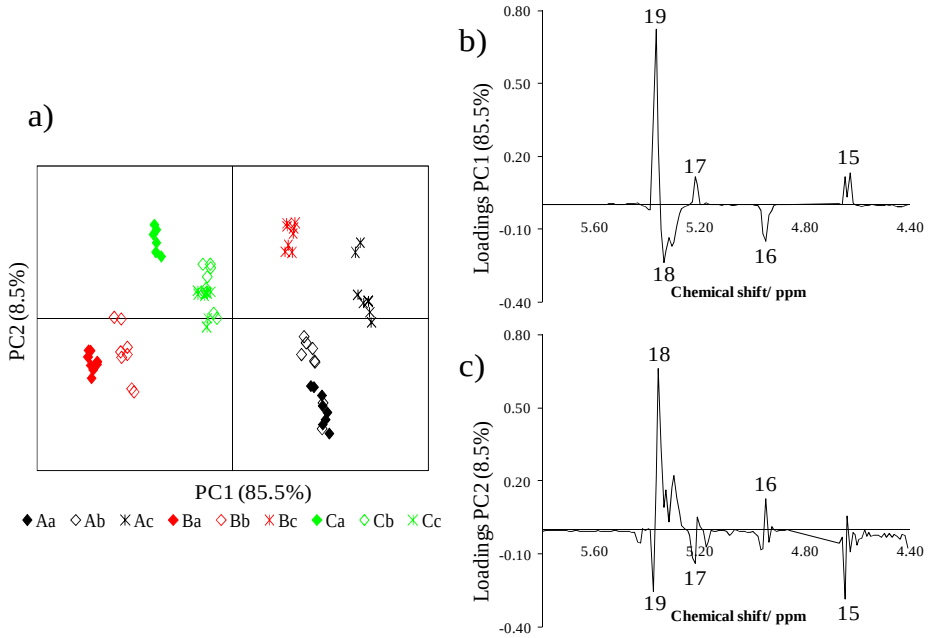

Figure 3.15. PCA for anomeric region (4.5-5.8 ppm) as a function of brewing sites and productions dates: a) PC1 vs. PC2 scores scatter plot, b) PC1 and c) PC2 loadings profiles. Signals varying are identified by the numbering shown in Figure 3.12.

Finally, for the aromatic region (5.8-10.0 ppm), significant differences are also observed in the separation between brewing sites when compared with the whole

spectra (Figure 3.16a). In this case, high data dispersion was observed for all the studied brewing sites (even between bottles from the same site and date), although a separation trend between sites is identified in the PCA scores plot, with PC1 and PC2 accounting for 68.7% and 17.0%, respectively, of the total variability. Based in the PCA results (Figure 3.16), it is suggested that beers from site **A** (black) and **C** (green, except site **Ca**), located in positive PC1, are distinguished from the beers produced in site **B** (red), in negative PC1, by higher amounts of acetaldehyde and aromatic amino acids (e.g. histidine, phenylalanine, tryptophan and tyrosine) and lower content in aromatic alcohols (i.e. 2-phenylethanol) and nucleosides. Moreover, beers from site **A**, with negative PC2, are separated from site **C** beers (except site **Ca**), located in positive PC2, reflecting higher amounts of acetaldehyde, formic acid and nucleosides as well as lower content in 2-phenylethanol and tryptophan. These variations (except for acetaldehyde) are in agreement with previous report (Almeida *et al*., 2006), where nucleosides, tyrosine and tyrosol (overlapped peaks), histidine and 2-phenylethanol show to vary within sites. Furthermore, comparing the compositional variations obtained for the aromatic region with those are considering the whole spectra (Figure 3.12) the same variations were noted, although, as mentioned above, a higher variability (even between bottles from the same site and date) and different separations between sites is observed.

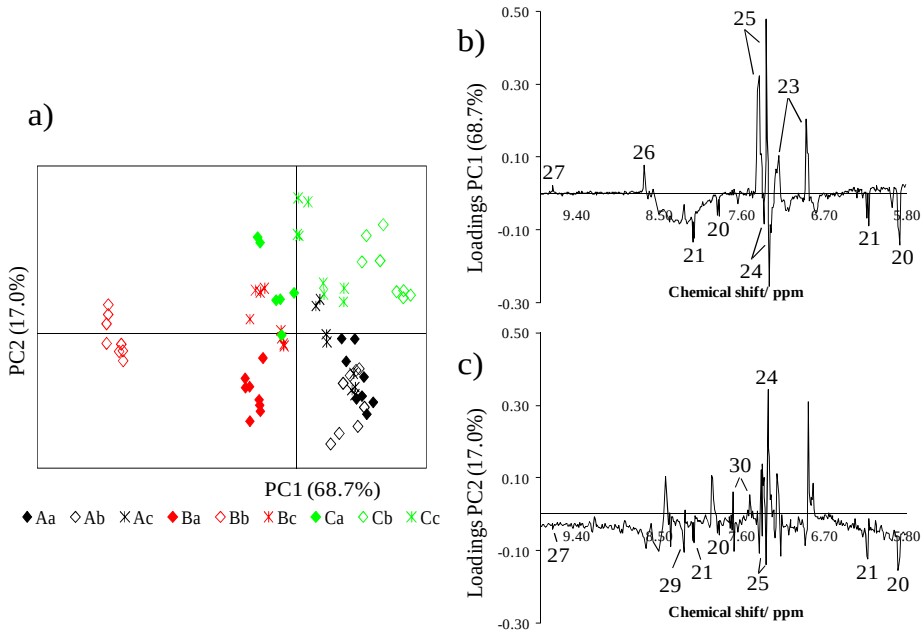

Figure 3.16. PCA for aromatic region (5.8-10.0 ppm) as a function of brewing sites and production dates: a) PC1 vs. PC2 scores scatter plot, b) PC1 and c) PC2 loadings profiles. Signals varying are identified by the numbering shown in Figure 3.11, with the addition of peaks 30, tryptophan.

Table 3.4. Compositional differences identified by NMR/PCA for lager beers (produced in three different brewing sites (**A**, **B** and **C**) and from three production dates (**a**, **b**, **c**). Amino acids are described by their three letter code. Aliphatic amino acids refers to: Ala, Arg, Ile, His, Leu, lLys, Val; aromatic amino acids: Phe and Tyr; higher alcohols: isobutanol, isopentanol and propanol; nucleosides: cytidine and uridine

Site	Reproducibility	Date	Compositional profile
A	High	**a**	**Richer (++):** aliphatic and aromatic amino acids/ linear dextrins/ acetaldehyde/ formic
		b	**High (+):** nucleosides
		c	**Low (-):** higher and aromatic alcohols/ branched dextrins/ citric/ lactic/ pyruvic/ succinic
B	Low	**a**	**Richer (++):**higher and aromatic alcohols/ branched dextrins/ nucleosides / citric/ lactic/ pyruvic / succinic **Low (-):** aliphatic and aromatic amino acids/ linear dextrin/ acetaldehyde
		b	**Richer (++):** higher alcohols/ citric/ lactic/ pyruvic / succinic/ branched dextrins/ aromatic alcohols/ nucleosides **Low (-)**:aliphatic amino acids/ linear dextrin/ / aromatic amino acids/ acetaldehyde
		c	**Richer (++):** aliphatic amino acids/ linear dextrins/ aromatic alcohols/ nucleosides **Low (-):** higher alcohols/ aromatic amino acids/ acetaldehyde citric/ lactic/ pyruvic/ succinic/
C	Medium	**a**	**Richer (++):** citric/ pyruvic/ succinic **High (+):** higher and aromatic alcohols/ aromatic amino acids/ branched dextrins/ acetaldehyde/ lactic **Low (-):** aliphatic amino acids/ linear dextrins/ nucleosides
		b	**Richer (++):**aliphatic amino acids / pyruvic **High (+):**aromatic alcohols/ aromatic amino acids/ branched dextrins/ acetaldehyde **Low (-):** higher alcohols/ linear dextrins/ nucleosides/ citric/ lactic/ pyruvic/ succinic
		c	**Richer (++):** higher alcohols / lactic **High (+):**aromatic alcohols/ aliphatic and aromatic amino acids/ branched dextrins/ acetaldehyde/ pyruvic **Low (-):** linear dextrins/ nucleosides

Assessment of PCA results was attempted by signal integration, so that variations could be confirmed or discarded. Overall, 26 compounds were integrated: alanine, GABA, leucine, histidine, isoleucine, proline, phenylalanine, tryptophan, tyrosine and valine (amino acids); higher alcohols (area comprising isobutanol, isopentanol and propanol signals); 2-phenylethanol and tyrosol (aromatic alcohols); acetic, citric, formic, gallic, lactic, malic, pyruvic and succinic acids (organic acids); adenosine/inosine, cytidine, uridine (nucleosides); and acetaldehyde. Examples of normalized areas for different brewing sites and production dates are shown in Figure 3.17 for a) alanine, b) tyrosine, c) higher alcohols, d) tyrosol, e) lactic acid, f) succinic

acid, g) cytidine, and h) acetaldehyde. The changes in dextrins content are also shown, being expressed in terms of variations in the average number of branching points/molecule and the average number of glucose units/molecule (Figures 3.17i,j, respectively).

In relation to amino acids, normalized area plots (examples for alanine and tyrosine in Figures 3.17a,b, respectively) show that beers from site **A** have higher reproducibility than beers produced in sites **B** and **C,** confirming the results obtained by PCA. Although similar quantities between the different sites and dates are observed, beers from **Bb** clearly show lower amounts of some amino acids (i.e. alanine, histidine, isoleucine, leucine, phenylalanine and tyrosine), possibly indicating a higher fermentation extent leading to a stronger uptake of amino acids.

For higher and aromatic alcohols, high compositional reproducibility is observed for sites **A** and **C**, whereas beers from site **B** show some variability between dates (especially for site **Bb**). Higher amounts of alcohols (higher alcohols and tyrosol shown in Figures 3.17c,d, respectively) are identified in beers from site **Bb (**opposite tendency that the one identified for amino acids). This reflects the lower fermentation extent and, consequently, weaker uptake of amino acids leading to a decreased formation of higher alcohols (by-products of amino acids assimilation, Ehrlich pathway). The characteristics of the yeast strain used, fermentation conditions (sluggish agitation, lower temperature and/or aeration deficiency), wort composition (low levels of nutrients such as amino acids, lipids, zinc and oxygen) and installation design are some of the factors that may lead to lower fermentation extent in brewing process (Ernandes *et al.*, 1993; Powell *et al.*, 2003; Willaert and Nedovic, 2006; Branyik *et al.*, 2008).

Organic acids variation within sites (lactic and succinic acids shown in Figure 3.17e,f, respectively) show some degree of variability, although for site **A** beers a better reproducibility is observed (confirming PCA information). Pyruvic and succinic acids (the latter seen in Figure 5.17f) and, to a less extent, citric acid show small quantities variations between sites and dates, with slightly lower amounts of those acids being observed in site **A** and in **Bc**. In addition, an interesting brewing process problem was identified in this study, with beers produced at site/date **Cc** showing significantly higher amounts of lactic acid (approximately 50 to 100% enhancement when compared with the other studied beer samples). This indicates a possible microbiological contamination by lactic acid bacteria (Bamforth, 2009) occurring in that brewing site.

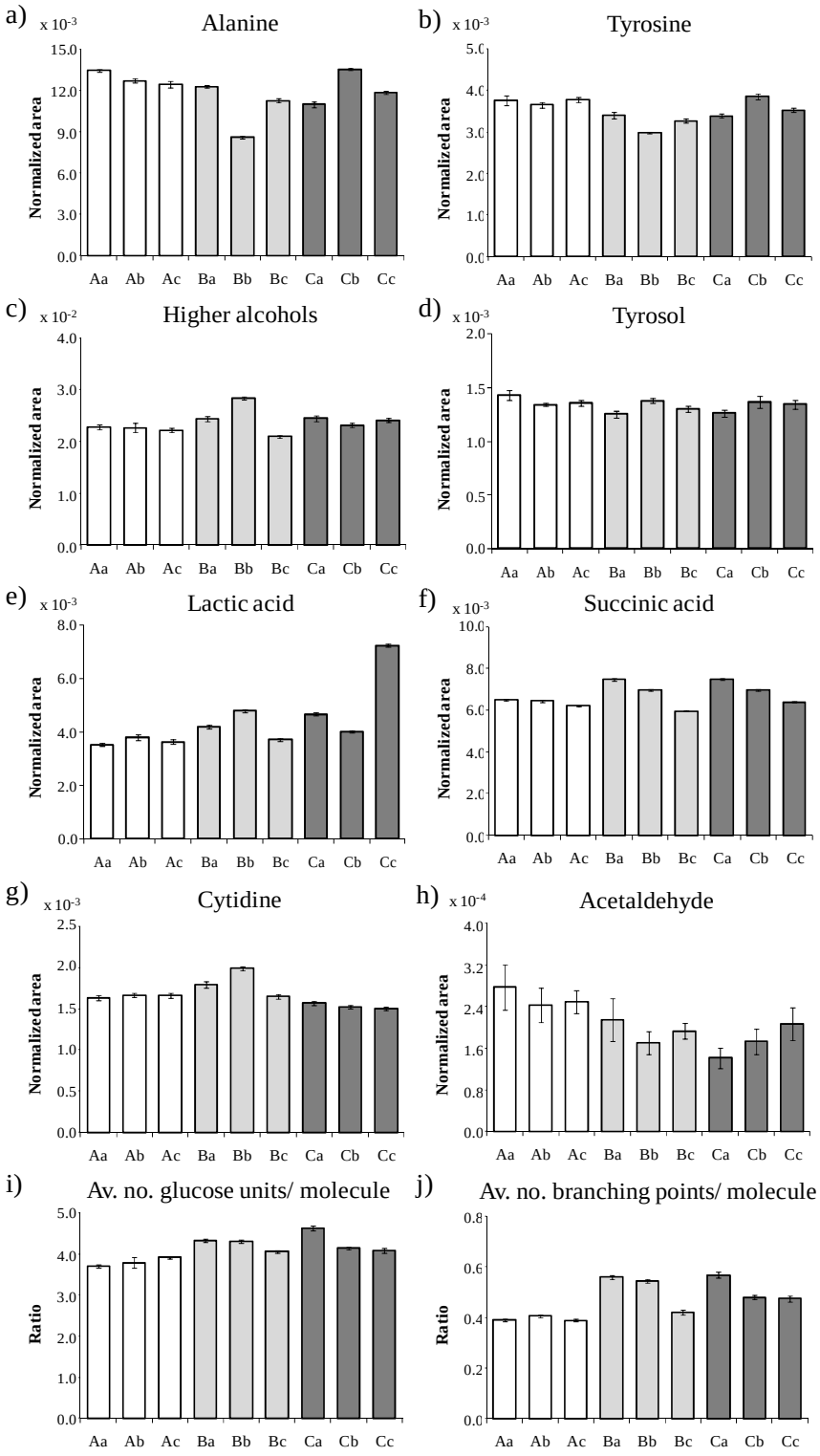

Figure3.17. Plots of normalized areas obtained for: a) alanine, b) tyrosine, c) higher alcohols, d) tyrosol, e) lactic, f) succinic acid, g) cytidine, and h) acetaldehyde, and of anomeric signals ratio reflecting g) average number of branching points per glucose oligomer molecule and h) average number of glucose units per glucose oligomer molecule, for different brewing sites and production dates. Beers from site **A** (white bars), site **B** (light grey bars), and site **C** (dark grey bars).

In relation to nucleosides, higher reproducibility is observed for sites **A** and **C**, when compared with those of site **B** beers, indicating a better process control in the sites **A** and **C** sites. Cytidine, uridine and, in less extent, adenosine/inosine (example shown for cytidine in Figure 5.17g) are present in slightly lower amounts in site **C** beers. This suggests the use of excessively high mashing temperatures, leading to a lower enzymatic degradation of nucleic acids and, consequently, lower amounts of nucleosides. In fact, nucleosides, together with nucleotides, sugars and free bases, are products of nucleic acids degradation occurring during malting and mashing (Lee and Prentice, 1987).

For acetaldehyde, some degree of variability within sites is seen for all brewing sites. In addition, high integration errors bars in the normalized area plot are observed, as a consequence of the low intensity of its peak (quartet, at 9.66 ppm) which affects the signals integration procedure. Higher quantities of acetaldehyde in beers from site **A** (Figure 3.17e) was confirmed, possibly indicating a sluggish fermentation step and/or lower fermentation temperature, decreased residence time of beer in the fermentation vessels, colder maturations and/or decreased yeast content during maturation. Furthermore, the use of poor quality pitching yeast, excessive wort oxygenation and pitching rates, and oxygen uptake of beer during maturation could also lead to an enhancement of the acetaldehyde content (Briggs, 2004). The control of acetaldehyde is extremely important in brewing industry, as it may originate a staling flavor in beer when present in excessive concentrations. Acetaldehyde is mainly produced during fermentation by yeast, being a by-product of carbohydrate metabolism as result of the decarboxylation of pyruvic acid (Priest and Stewart, 2006; Branyik *et al.*, 2008; Jonkova and Petkova, 2011).

Regarding dextrins content (corresponding to average number of glucose units and average number of branching points per molecule shown in Figures 3.16g,f, respectively), beers from site A show a higher reproducibility within sites. In addition, beers produced in sites **A** and **Bc** (the latter in less extent) show higher amounts of linear dextrins. This higher content in linear oligomers may indicate the use of milder mashing conditions (in terms of time and temperature), influencing the ratio of α- and β-amylase activities (Uchida *et al.*, 1991), and the activities of the heat sensitive α-1,6 debranching enzymes. Dextrins are unfermentable oligomers, therefore, the ratio

linear/branched dextrins should reflect the relation between the activities of both α- and β-amylase occurring during the mashing process.

In summary, beers from site **A** show a good overall compositional reproducibility within dates, although in the aromatic region some slight dispersion is detected. For site **B** beers, significant dispersion of beers samples in all the studied spectral regions was noted, particularly in beers from **Bc** in aliphatic and aromatic regions and **Bb** in anomeric region. Finally, for site **C** beers, although a good reproducibility is noted in the anomeric and, in less extent, in aromatic regions, significant samples dispersion in the aliphatic region occurs. Overall, site **A** beers show an overall higher reproducibility, suggesting a better control of the brewing process parameters.

Regarding beer composition, beers from site **A** show higher amounts of amino acids, acetaldehyde and linear dextrins, whereas lower quantities of higher alcohols, citric, pyruvic and succinic acids were also identified. These increases of amino acids and decreases of higher alcohols and the mentioned acids (also observed for beers from site/date **Bc**) indicates a lower fermentation extent occurring in site **A**. The higher amounts of acetaldehyde shown for beers from site **A** are in agreement with the low fermentation extent possibility. Furthermore, beers from site **A** also show higher amounts of linear oligomers indicating milder mashing conditions in the production process.

As mentioned before, site **B** beers show significant compositional variability within production dates. For example, beers from site **Bb** show lower amino acids and acetaldehyde content together with higher amounts of higher alcohols, indicating a higher fermentation extent.

Beers from site **C** are characterized by slightly lower amounts of cytidine, uridine and, in less extent, adenosine/inosine suggesting the use of excessively high mashing temperatures. Furthermore, significantly higher amounts of lactic acid are present in beers from site **Cc** (possible microbiological contamination by lactic acid bacteria occurring in that brewing site).

3.6 Conclusions

In this chapter, a detailed chemical characterization of a lager beer was performed by NMR methods. By applying 1D and 2D NMR and LC-NMR combined with LC-MS, 40 metabolites were assigned in beer, including 15 amino acids (Ala, GABA, Arg, Glu, Gln, His, Ile, Leu, Lys, Phe, Pro, Thr, Trp, Tyr and Val), 8 acids (acetic, citric, formic, gallic, lactic, malic, pyruvic and succinic acids), 7 alcohols (ethanol, glycerol, isobutanol, isopentanol, 2-phenylethanol, propanol and tyrosol), 4 nucleosides (adenosine, cytidine, inosine and uridine), 2 sugars (maltose and trehalose), acetaldehyde, 5-HMF and the tentative assignment of acetoin and ethyl acetate, as well as dextrins resonances. Although no further compound assignment was achieved when compared to literature, this study enabled the detailed characterization of beer using pH adjusted to acidic values, being this information crucial for the following metabonomic studies performed in this thesis.

In the second part of this chapter, the NMR/MVA approach was applied to ^{1}H NMR data of a set of lager beers from the same brand, produced in three different brewing sites (located in the same country, Portugal) and three different dates, for evaluating beer variability. This study enabled the assessment of fine compositional variations within sites and dates and, subsequently, the brewing process performance and their effects on the final product composition since all materials and processes are uniform, being unpredicted shifts to occur particularly at the processing level. In fact,by NMR/MVA and signals integration, it was clearly observed that beers produced in site **A** show the highest reproducibility within dates, indicating a well-controlled brewing process and, consequently, the production of more reproducible beers. On the other hand, site **B** beers show great variability, indicating a lower control of the process in that site. In terms of beers composition and their relation with the production process, NMR/MVA approach enabled the identification of several process-related compositional variations. For instance, beers from site **A** (also observed for beers from site/date **Bc**) showed higher amounts of amino acids and lower quantities of higher alcohols, citric, pyruvic and succinic acids, which indicates the occurrence of a lower fermentation extent in site **A**, possibly due to the fermentation conditions used (e.g. sluggish agitation, lower temperature and/or aeration deficiency). Acetaldehyde is also present in higher amounts in site **A** beers, possibly reflecting the mentioned less fermentation extent and/or low efficiency of the maturation step. Slightly lower

quantities of nucleosides in site **C** beers indicate the use of excessively high mashing temperature. Finally, beers from site **A** also show higher amounts of linear oligomers, indicating milder mashing conditions (in terms of time and temperature) in the production process. In addition, a potential microbiological contamination by lactic acid bacteria was detected, indicated by excessive amounts of lactic acid present in beers produced in site **Cc**, thus demonstrating the applicability of this approach in the on-line control of brewing process.

Overall, the NMR/MVA methodology was demonstrated to be a useful tool for the short-term monitoring of brewing process, enabling the identification of finer compositional deviations related to slight difference in the production process, even in well-controlled brewing processes as the ones studied. These results unveil the possibility for a future routine application of this methodology in the control of beer production.

4. QUANTIFICATION OF ORGANIC ACIDS IN BEER BY NMR METHODS: A CRITICAL COMPARISON STUDY

4.1 Introduction

The quantification of organic acids in beer is of special interest to the brewing industry as these compounds reflect characteristics related to wort composition and the brewing process (e.g. fermentation extent) as well as of the final product, since they influence the organoleptic properties of beer and its final pH (Belke and Irwin, 1992; Boulton and Quain, 2001; Hughes and Baxter, 2001). The determination of organic acids in beer has been carried out mainly by a variety of chromatographic methods (Yan *et al.*, 1997; Perez-Ruiz *et al.*, 2004), however time-consuming and, in most cases, needing complex sample preparation, and by an increasingly use of capillary electrophoresis (CE) (Klampfl, 1999; Cortacero-Ramirez *et al.*, 2005; Nord *et al.*, 2004; Erny *et al.*, 2009).

NMR spectroscopy is presented as an attractive alternative method for organic acid quantification in beer, as it enables direct sample analysis and can be applied in tandem with automation and multivariate analysis methods in order to handle large sample numbers, showing high accuracy. However, although NMR is an inherently quantitative technique, important stumbling blocks arise when direct quantification in complex matrices, as in beer, is attempted. The traditional method of integration vs. the area of a reference compound has been widely applied for the quantification of specific compounds (e.g. amino acids, carbohydrates, organic acids and alcohols), with applications in vinegars (Consonni and Gatti, 2004; Caligiani *et al.*, 2007), wines (Viggiani and Morelli, 2008; Lopez-Rituerto *et al.*, 2009) and fruit juices (Berregi *et al.*, 2003; del Campo *et al.*, 2006; Spraul *et al.*, 2009) as well as in beer (Nord *et al.*, 2004; Lachenmeier *et al.*, 2005). However, the use of internal references poses potential difficulties as possible signal overlap and chemical interactions between the reference compound and sample components may occur, leading to changes in signal area and/or in shape and, consequently, to erroneous quantification. An alternative is the use of the Electronic Reference To access In vivo Concentrations (ERETIC) method, based on an electronic reference signal and thus improving quantification precision (Barantin *et al.*, 1997; Akoka and Trierweiller, 2002; Rizzo and Pinciroli, 2005). Reports of the use of this method have been described for food extracts, i.e. for carrots roots (Cazor *et al.*, 2006) and tomato fruit extracts (Deborde *et al.*, 2009). It is important to note that, whatever the signal integration method chosen, NMR analysis requires quantitative acquisition conditions, which should allow complete T_1 relaxation of all spins.

Therefore, the use of inter-scan times of at least 5 x highest T_1 is necessary for the accurate quantification of all compounds under study (Evilia, 2001), increasing analysis times and making this approach time-consuming.

Quantitative multivariate methods, namely PLS1 regression, have been increasingly considered as a suitable method for the rapid quantification of compounds, allowing a routine application to large sample populations. The PLS1 regression model is obtained by maximizing the covariance between the NMR spectra with a calibration set of concentrations measured by a reference analytical method. This procedure, already used in wine (Larsen *et al.*, 2006), fruit juices (Spraul *et al.*, 2009) and beer (Nord *et al.*, 2004; Lachenmeier *et al.*, 2005), enables the use of shorter NMR runs through semi-quantitative spectra, circumventing the difficulties associated with time-consume, and tackles errors associated to signal integration.

In this chapter, a comparative study of the quantification of organic acids in beer using different NMR methods is described. The determination of organic acids in beer have been addressed in previous reports, either by using NMR integrative methods (Nord *et al*, 2004) as well as by quantitative MVA methods, PLS1 regression, using as calibration reference capillary electrophoresis (CE) with indirect detection (Nord *et al.*, 2004) and enzymatic essays (Lachenmeier *et al.*, 2005) results. In this thesis, different quantitative NMR methodologies were developed and compared. In this way, ^{1}H NMR data was used to quantify the six main organic acids in beer, namely acetic, citric, lactic, malic, pyruvic and succinic acids, by 1) PLS1 regression models using as quantitative reference method CE with indirect detection, CE with direct detection and enzymatic essays; and 2) NMR integration both vs. an internal reference compound and using the ERETIC method.

4.2 Experimental section

4.2.1 Beer samples and sample preparation

All lager beer samples (a total of 95) were of the same brand and were kindly donated by UNICER, Bebidas de Portugal. Several groups of samples, differing in date of production, were employed for the development of multivariate models (calibration sets) and for their testing (test sets), as shown in Table 4.1. The bottled samples were kept unopened at 4^{o}C, for a period of up to 15 days prior to analysis.

Table 4.1. Beer sample sets employed as calibration and test sets.

Beer sample groups / analytical method	No. beer production dates	No. bottles/date	Total no. samples
Calibration sets			
Group A / CE with indirect detection	3	12	36
Group B / CE with direct detection	3	9	27
Group C / Enzymatic essays	6	3	18
Test sets			
Group D	1	6	6
Group E	1	8	8

Beer samples (10 mL) were degassed in an ultrasonic bath for 10 minutes. For NMR analysis, samples were prepared to contain 10 % D_2O and 0.025 % TSP-d_4 as chemical shift and intensity reference. Samples pH was adjusted to 1.90 ± 0.03 adding 8-10 µl HCl 5% in D_2O, in order to ensure extensive protonation of carboxylic groups. For CE analysis, samples were diluted 3 times and filtered through a 0.22 µm polyvinylidene fluoride (PDVF) membrane filter. For enzymatic analysis, the degassed beer was used without any preparation.

4.2.2 Chemicals

All chemicals were of analytical grade: 3-(trimethylsilyl)propionic-2,2,3,3-d_4 acid, sodium salt (TSP-d_4), glyoxilic monohydrate, malic and pyruvic acids (Aldrich), succinic acid (Carlo Erba), tetradecyltrimethylammonium bromide (TTAB), citric and lactic acids (Sigma), sodium dihydrogen phosphate dihydrate and disodium hydrogen phosphate (Fluka), 2,6-pyridinedicarboxylic acid (PDC), calcium chloride dihydrate, potassium dihydrogen phosphate, sodium hydroxide, hydrochloric acid and acetic acid (Merck), cethyltrimethylammonium bromide (CTAB), ethylenediaminetetraacetic acid (EDTA) (Panreac). Tetradecyltrimethylammonium hydroxide (TTAOH) was prepared from the bromide salt (TTAB) using a strong anion-exchange resin AG MP-1 (Biorad).

4.2.3 NMR spectroscopy

T_1 experiments on a lager beer were recorded on a Bruker Avance DRX 500 spectrometer, at 300 K, in order to determine the correct T_1 relaxation times of the studied organic acids. The T_1 experiments were measured using *t1zgpr* pulse recovery sequence (180^o_x-τ-90^o_x), with delays between pulses (τ) in 100ms – 30s range (30 experimental points).

For PLS1-NMR regression studies, each spectrum was acquired on a Bruker Bruker Avance DRX 500 spectrometer equipped with a 5 mm inverse probe, at 300 K, using the *noesypr1d* pulse sequence with water presaturation. 128 transients were collected into 32768 (32k) data points, with mixing time of 100 ms, spectral width of 8013 Hz, acquisition time of 2.0 s and relaxation delay of 5.0 s.

For quantification by integration vs. TSP or by ERETIC, quantitative conditions (i.e. relaxation delay ≥ 5 x longest spin-lattice relaxation time ($T_{1longest}$) corresponding to organic acids) were employed based on the measured T_1 relaxation times (3.8 s acetic, 1.1 s citric, 1.3 s lactic, 1.5 s malic, 2.9 s pyruvic and 1.7 s for succinic acids). Spectra for integration vs. TSP were recorded on a Bruker Avance DRX 500 spectrometer, at 300 K, using the *zgpr* pulse sequence (single 90º pulse experiment with water suppression), with 128 transients collected into 32768 (32k) data points, spectral width of 8013 Hz, acquisition time of 2.0 s and relaxation delay of 20.0 s. ERETIC measurements were recorded on a Bruker Avance III 600 spectrometer, at 300 K, using the same *zgpr* pulse sequence, with 128 transients collected into 65536 (64k) data points, spectral width of 12336 Hz, acquisition time of 2.7 s and relaxation delay of 25.0 s. A standard succinic acid solution (119.90 mg L^{-1}) was used to calibrate the intensity of ERETIC electronic signal. Each FID was zero-filled to 64 k points and multiplied by a 0.3 Hz exponential line-broadening function prior to FT. All spectra were manually baseline corrected and chemical shifts referenced internally to the TSP peak or to the electronic signal of ERETIC. All peak assignments were carried out as described in section 3.3.

Two ^{1}H NMR quantification methods based in signals integration were used: vs. the area of a reference compound and by ERETIC. In NMR integration vs. the area of a reference compound, i.e. TSP signal, the concentration of each compound (in mgL^{-1}), m_X, was calculated as (Caligiani *et al.*, 2007; Viggiani and Morelli, 2008):

$$m_X = (A_X / A_{TSP}) * (m_{TSP} / (MW_{TSP} / no.H_{TSP})) * (MW_X / no.H_X) \qquad [Eq. 4.1]$$

where m_{TSP} refers to concentration of TSP, A_X and A_{TSP}: peak areas, MW_X and MW_{TSP}: molecular weights and $no.H_X$ and $no.H_{TSP:}$ number of hydrogens corresponding to the peaks from compound X and TSP.

For quantification by ERETIC method (Akoka *et al.*, 1999; Akoka and Trierweiler, 2002), calibration of the electronic reference signal was carried out using a

succinic acid standard solution (119.90 mg L^{-1}) and the methylene signal at 2.66 ppm. The equivalent concentration of the electronic signal, m_{ERETIC}, was determined by:

$$m_{ERETIC} = (A_{ERETIC} / A_{REF}) * m_{REF} \quad \text{[Eq. 4.2]}$$

where A_{ERETIC} and A_{REF} are the areas of the electronic signal and calibration peak (succinic acid), respectively. The electronic signal was then used to determine the concentration of analyte *X*, m_X, by computing:

$$m_X = (A_X / A_{ERETIC}) * (m_{ERETIC} / no.H_X) * (MW_X / MW_{REF}) \quad \text{[Eq. 4.3]}$$

where A_X: peak area for the analyte, $no.H_X$: number of hydrogens corresponding to the analyte signal and MW_X and MW_{REF}: molecular weight of analyte and calibration compound (succinic acid), respectively.

For the quantitative analysis by NMR integration, the selected peaks integration was carried out by manually picking start and stop positions in the spectra with AMIX.3.9.5, Bruker BioSpin. Peaks showing lower overlap (based on observation of both 1D and 2D spectra, shown in Figures 3.3 and 3.4) were chosen for integration: acetic at 2.08 ppm, citric at 3.01 ppm, lactic at 1.41 ppm, malic at 2.89 ppm, pyruvic and hydrate form at 2.40 and 1.56 ppm, respectively, and succinic at 2.66 ppm. Some overlapping contributions were noted, namely from i) proline (2.02 and 2.37 ppm) and acetic (2.08 ppm) and pyruvic (2.40 ppm) acids, ii) GABA (3.02 ppm) and citric acid (3.01 ppm), iii) unassigned signal at 2.85 ppm and malic acid (2.89 ppm) and iv) isopentanol (1.40 ppm) and lactic acid (1.41 ppm), although in the latter case in lesser extent.

4.2.4 Capillary Electrophoresis (CE) analysis

Capillary zone electrophoresis method was employed. Two types of UV detection modes were used: direct and indirect detection.

All experiments were performed using a Beckman P/ACE MDQ CE system, equipped with a diode array detector. Standards and samples were injected using 0.3 psi (2068 Pa) pressure for 2 seconds. Separations were carried out in a fused silica capillary of 78 cm total length and 75 μm internal diameter, at a potential of -25 kV and -17 kV (reverted polarity) for direct and indirect UV detection, respectively, maintaining a constant capillary temperature of 18 ºC. For indirect detection, a method previously developed for the analysis of organic acids in port wine (Esteves *et al.*, 2004) was used,

employing the 196 nm wavelength, together with reference wavelength at 350 nm. The run buffer used was composed of 5 mM of PDC and CTAB 0.5 mM, acting as electroosmotic modifier, with pH adjusted to 5.60 ± 0.05, with 1M NaOH. In order to eliminate interferences from trace metals on the determination of citric acid, 0.01 mM of EDTA was added to the buffer. For direct detection, the method reported in (Mato *et al.*, 2007) was used, employing a wavelength of 195 nm. The run buffer was composed of 7.5 mM NaH_2PO_4 and 2.5 mM Na_2HPO_4 with TTAOH 2.5 mM and $CaCl_2$ 0.24 mM being used as electroosmotic modifier and selectively modifier, respectively, and pH adjustment to 6.4 ± 0.05, with 1M NaOH. Glyoxilic acid was added as internal standard although it was eventually not used, due to peak overlap with contributions from the beer matrix. Peaks integration was carried out with P/ACE MDQ CE software.

4.2.5 Enzymatic methods

The reagents for the enzymatic essays used were purchased from Megazyme. The essays were based on an increase/decrease in absorbance at 340 nm, caused by a change in nicotinamide adenine dinucleotide (reduced form), NAD(H). Absorbances were measured in an ASYS UVM spectrometer reader using 96 well microplates and results were affected by an average associated error of 8%.

4.2.6 Regression models

For PLS1 regression analysis, data matrices **X** were built using the aliphatic region from 1.35 to 3.10 ppm. Spectral bucketing was employed with variable bucket width (to minimize the 1st derivative-like effects), using Amix-Viewer 3.9.2 software. Capillary electrophoresis, using both direct and indirect ultraviolet (UV) detection, and enzymatic essays were used as reference methods (the **y** data vector). All PLS1 regression studies were performed using an in-house application co-developed by the University of Aveiro and the AgroParis Tech, France. In order to facilitate and improve the interpretation of the PLS1 regression models, the loadings weights profiles were colored as a function of the correlation between each NMR data point and the organic acid content, as expressed in the **y** data vector, being the colored loadings plots obtained using R-statistical software 2.11.0 and Plotrix package (Lemon, 2010).

4.3 Optimization of reference methods for the quantification of organic acids

For all organic acids studied (acetic, citric, lactic, malic, pyruvic and succinic acids), calibration curves were built for each used analytical method: CE with direct detection, CE with indirect detection and enzymatic essays. For each organic acid, the concentration ranges used in the calibration curves were chosen based on concentration values found in literature for organic acids in beer (Hardwick, 1994; Klampfl, 1999; Montanari *et al.*, 1999; Hughes and Baxter, 2001). The results obtained are listed in Table 4.2.

Table 4.2. Results obtained for the calibration curves obtained for CE with direct detection, CE with indirect detection and enzymatic essays methods, used in the quantification of organic acids in beer.

	CE (direct detection		**CE (indirect detection**		**Enzymatic assays**		
Organic acids	Linear range (mgL^{-1}) [a]	r [a]	Linear range (mgL^{-1}) [a]	r [a]	Linear range (mgL^{-1}) [a]	r [a]	Expected conc. (mgL^{-1}) [b]
Acetic	20.1 – 140.1	0.97	39.9 – 159.9	0.79	20.0 – 180.0	0.98	44-150 [c] 37-171 [d]
Citric	---	---	30.0 – 300.0	0.99	55.0 – 274.8	0.99	--- 56-230 [d]
Lactic	25.5 – 177.6	0.98	25.2 – 250.8	0.98	20.7 – 206.6	0.99	65-118 [c] 44-292 [d]
Malic	20.4 – 141.9	0.98	20.1 – 201.0	0.99	40.6 – 202.9	0.98	45-103 [c] 14-105 [d]
Pyruvic	15.0 – 105.0	0.98	15.0 – 150.0	0.96	5.0 – 35.0	0.98	28-93 [c] 10-104 [d]
Succinic	15.3 – 106.5	0.97	15.0 – 150.0	0.98	30.4 – 121.5	0.95	50-108 [c] 36-166 [d]

[a] Linearity range used and correlation coefficients obtained for the calibration curves of each reference method. [b] Expected concentration range in beer: values reported in [c] (Hardwick, 1994) and [d] (Montanari *et al.*, 1999).

Figure 4.1 shows the CE electropherograms obtained for a standard mixture of the six organic acids (Figure 4.2a) and a beer sample (Figure 4.2b), using both direct and indirect detection (respectively, left and right in Figure 4.1). All assignments were confirmed using standard solutions of each acid. It is clear that the mode of detection significantly determines the quality of the electropherograms, particularly affecting citric acid detection (peak 2). In fact, a broad peak is obtained for citric acid in direct detection (Figure 4.1a, left) whereas, if indirect detection is employed, the same acid

shows up clearly at a different separation time (difference in migration times between the two electropherograms is due to differences in buffer pH and ionic strength).

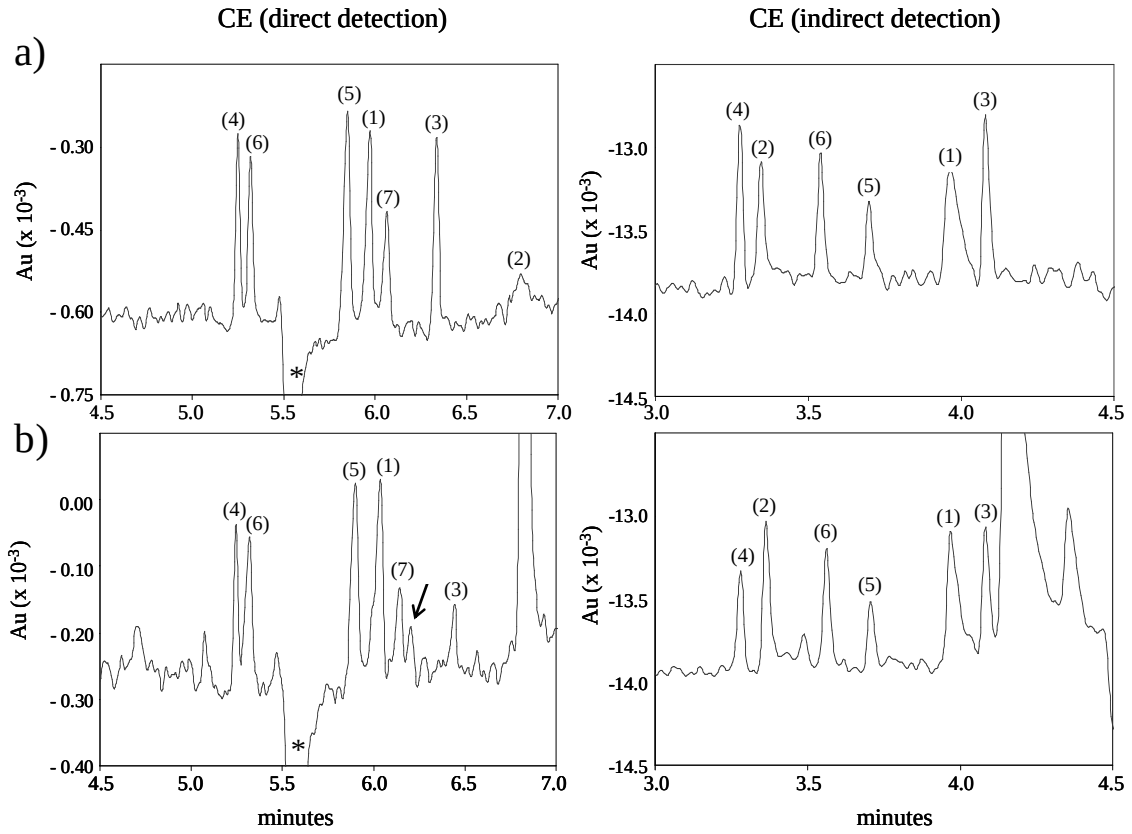

Figure 4.1. CE electropherograms of a) a standard organic acids mixture and b) beer (diluted of 1:3) recorded with direct (left) and indirect (right) detection. (1) acetic, (2) citric, (3) lactic, (4) malic, (5) pyruvic, (6) succinic and (7) glyoxylic acids (internal standard). * system peak.

Indirect detection seems to suit most organic acids in the standard mixture, with only some broadening being noted for peak 1 (acetic acid). Consistently, the CE calibration curves obtained for direct detection (Table 4.2) give correlation coefficients $r > 0.97$ for all acids, except for citric acid, which could not be quantified applying this method. For indirect detection, calibration curves are characterized by $r > 0.96$ with the exception of acetic acid ($r = 0.79$), reflecting the peak broadening noted before (Figure 4.1a). When beer samples are considered (Figure 4.1b), additional (unidentified) peaks arising from the beer matrix hinder citric acid detection further (peak 2) and overlap significantly with peak 7 (reference compound glyoxylic acid, marked with an arrow, thus preventing its use as internal standard) in direct mode, while affecting peaks 1 and 3 (acetic and lactic acids) in indirect mode. Furthermore, some peaks are clearly

broadened in direct detection mode: peaks 1 (acetic acid) and 6 (succinic acid), the latter affecting the neighbouring peak 4 (malic acid).

Enzymatic essays were also used to quantify all six organic acids and good linearity for the corresponding calibration curves is observed, with correlation coefficients, r > 0.95 (Table 4.2), although problems in terms of reproducibility are observed.

4.4 Development of PLS1-NMR regression models and comparison with NMR integrative methods

Figure 4.2 shows a typical spectrum of a lager beer, dominated by resonances arising from dextrins, in the mid field region, and from organic acids and amino acids in the remaining regions.

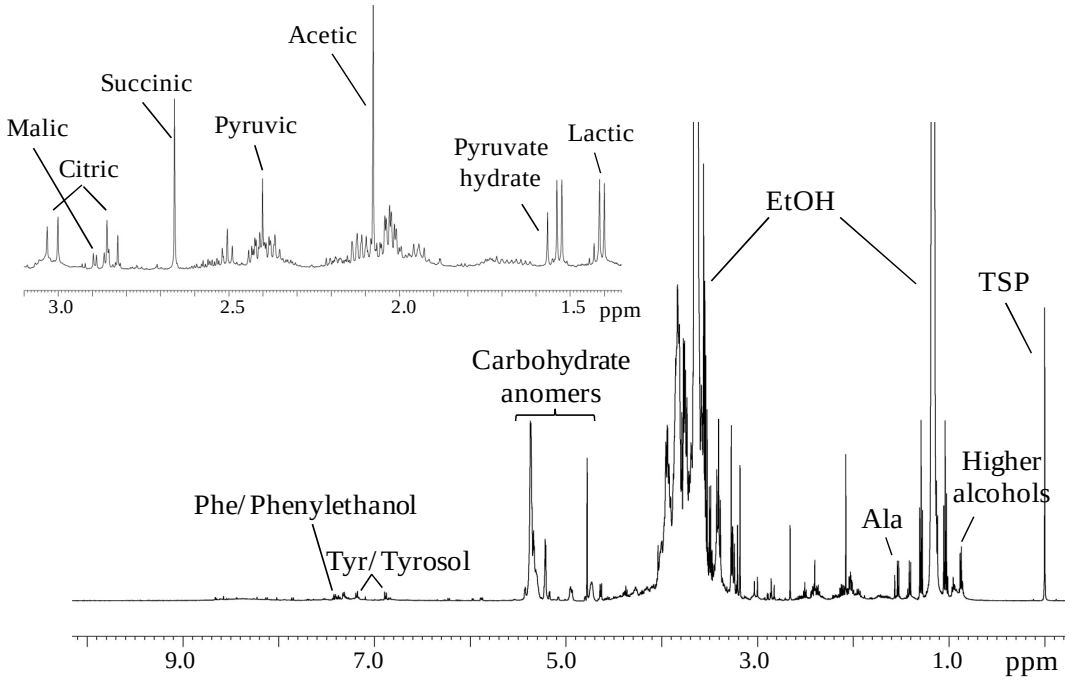

Figure 4.2. 500 MHz ^{1}H NMR spectrum of beer with some assignments indicated. The inset shows the expansion of the aliphatic spectral region (1.35-3-10 ppm) dominated by the organic acids signals, together with higher alcohols (isobutanol, isopentanol and propanol) and amino acids (e.g. alanine, GABA and proline).

The expansion of the 1.35-3.10 ppm region clearly shows resonances from all six main organic acids to be quantified in this work: acetic, citric, lactic, malic, pyruvic and succinic. Therefore, this region was chosen for the built up of all PLS1-NMR regression

models as well as for the integration of selected peaks for each organic acid. PLS1-NMR regression models, using as reference the response of CE (with direct and indirect detection) and enzymatic essays, developed for beer were found to correspond to numbers of Latent Variables (LVs) ranging from 3 to 8, as seen in Table 4.3.

Table 4.3. Results obtained for the built PLS1-NMR regression models for the quantification of organic acids in beer.

Organic acids	δ (ppm) [a]	PLS1-NMR regression models results				
		LV [b]	R^2Y (%) [c]	RMSECV (%) [d]	RMSEP (%) [e]	R [f]
Capillary Electrophoresis (direct detection)						
Acetic	2.08, s, CH_3	3	94.0	5.5	8.8	0.97
Citric	2.84, d, CH	---	---	---	---	---
	3.01, d, CH					
Lactic†	1.41, d, CH_3	5	91.8	8.5	20.3	0.96
Malic†	2.89, dd, CH	6	72.6	6.5	16.5	0.85
	2.92, dd, CH					
Pyruvic	2.40, s, CH_3	4	92.1	4.6	10.0	0.96
Succinic†	2.66, s, CH_2	8	76.5	5.5	18.9	0.87
Capillary Electrophoresis (indirect detection)						
Acetic†	2.08, s, CH_3	8	93.0	9.7	25.4	0.96
Citric	2.84, d, CH	8	83.5	4.1	9.9	0.91
	3.01, d, CH					
Lactic†	1.41, d, CH_3	5	78.3	17.9	32.1	0.89
Malic	2.89, dd, CH	8	92.2	2.8	8.9	0.96
	2.92, dd, CH					
Pyruvic	2.40, s, CH_3	8	96.0	5.1	13.2	0.98
Succinic	2.66, s, CH_2	6	91.6	2.7	4.7	0.96
Enzymatic essays						
Acetic	2.08, s, CH_3	4	88.2	5.4	10.6	0.91
Citric	2.84, d, CH	6	92.2	2.6	7.6	0.96
	3.01, d, CH					
Lactic†	1.41, d, CH_3	4	76.1	12.1	24.9	0.87
Malic	2.89, dd, CH	6	91.2	6.8	25.2	0.96
	2.92, dd, CH					
Pyruvic	2.40, s, CH_3	6	87.5	3.9	11.4	0.94
Succinic	2.66, s, CH_2	4	82.4	8.8	16.6	0.91

[a] Assignments refer to resonances in the spectral region considered: 1.35-3.10 ppm; multiplicity (s, singlet; d, doublet; dd, doublet of doublets). [b] Number of latent variables obtained by cross-validation. [c] Explained accumulated variance of **y** data for each PLS1-NMR model. [d] Root Mean Square Error of Cross Validation. [e] Root Mean Square Error of Prediction. [f] Correlation coefficients corresponding to the PLS1-NMR models. † Acids for which the PLS1-NMR performance is low.

The models using CE/direct detection as the reference method were characterized by accumulated R^2Y ranges from 72.6 to 94.0% for all acids (excluding citric acid - not quantified by this method), with lower values (< 7%) for malic and succinic acids (probably reflecting the broadening of the corresponding CE peaks, as noted above). Hence, the prediction power of these PLS1 regression models is relatively low for these acids as expressed by the high values (>15%) of the root mean square error of prediction (RMSEP). A similar result is also noted for lactic acid (RMSEP > 20%), probably reflecting the fact that the lower intensity of the lactic acid peak in beer (peak 3 in Figure 4.1) leaves it rather close to the noise level and, therefore, more prone to quantification error.

The PLS1-NMR models obtained with CE/indirect detection (Figure 4.3) exhibit reasonable correlation coefficients (> 0.89) for all acids, however, a lower accumulative R^2Y value (78.3 %) is found for lactic acid, with RMSEP > 25% noted for both lactic and acetic acids. These findings could reflect the peak broadening and overlap effects noted previously for these two acids, in indirect detection mode (Figure 4.1).

Finally, for the PLS1-NMR model obtained using enzymatic method as reference (Figure 4.4, Table 4.3), good general prediction powers are found, again with the exception of lactic acid (with low accumulative R^2Y and high % RMSEP) and malic acid (with RMSEP > 25%). In the case of malic acid, a cluster of samples is seen at higher concentrations and, if samples are removed, the model is indeed significantly improved (RMSEP 14.7%). However, lager beers may contain malic acid in a wide range of concentrations (Klampfl, 1999; Nord *et al*, 2004), thus justifying a model built on the concentration range shown in Figure 4.5. The organic acids identified with the symbol † in Table 4.3 are those for which lower PLS1-NMR predictive power is found, thus showing that the PLS1-NMR regression models based on CE indirect detection and enzymatic methods are the most suitable (for 4 acids out of the total 6 for both methods), in spite of none having a satisfactory predictive power for lactic acid. In the future, a possible improvement for these PLS1 regression methods is the use of a larger universe of samples in order to have more robust quantitative models.

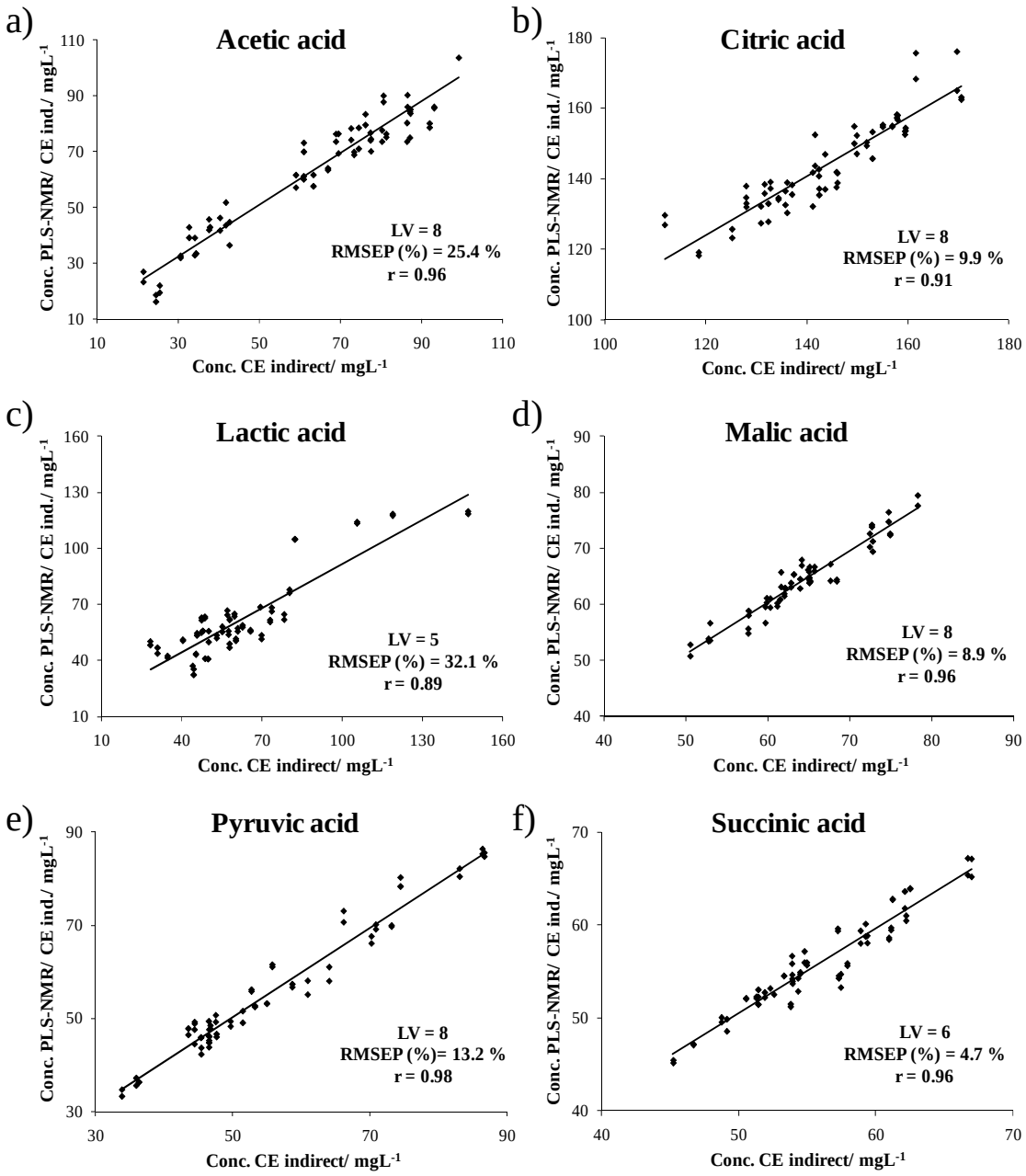

Figure 4.3. PLS1-NMR prediction models using CE (indirect detection) as reference for a) acetic, b) citric, c) lactic, d) malic, e) pyruvic and f) succinic acids. The number of latent variables (LV), RMSEP (%) and correlation coefficients (r) are shown for each model.

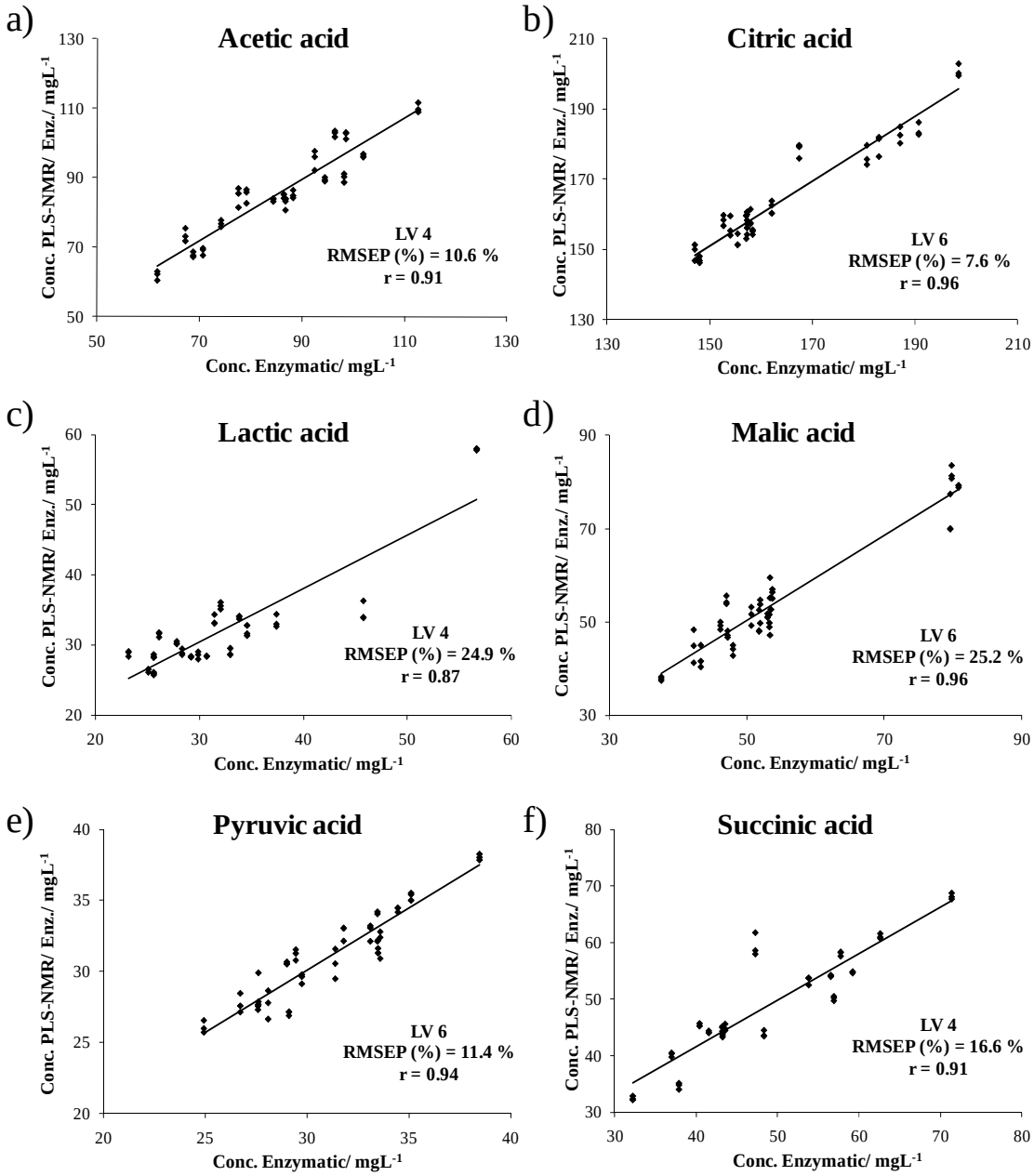

Figure 4.4. PLS1-NMR prediction models using enzymatic essays as reference for a) acetic, b) citric, c) lactic, d) malic, e) pyruvic and f) succinic acids. The number of latent variables (LV), RMSEP (%) and correlation coefficients (r) are shown for each model.

In order to facilitate the interpretation of the variability contained in each PLS-NMR regression model, Figure 4.5 shows examples of LV1 loadings weights (**w1**) profiles obtained for some of the PLS1-NMR regression models. The profiles are colored as a function of the correlation between each NMR data point and the **y** vector data and they show that maximum correlations (in red) are indeed obtained for the peaks arising from the organic acid under study in each case: acetic and lactic acid by PLS1-

NMR with CE/direct detection (Figure 4.5a,b) and pyruvic acid by PLS1-NMR with CE/indirect detection (Figure 4.5c). However, in all cases, additional peaks arising from some of the remaining acids are noted with non-zero correlations. These observations should reflect concomitant variations within the different organic acids e.g. in beer with higher acetic acid content, succinic, pyruvic and lactic acids seem to tend to lower concentrations (Figure 4.5a).

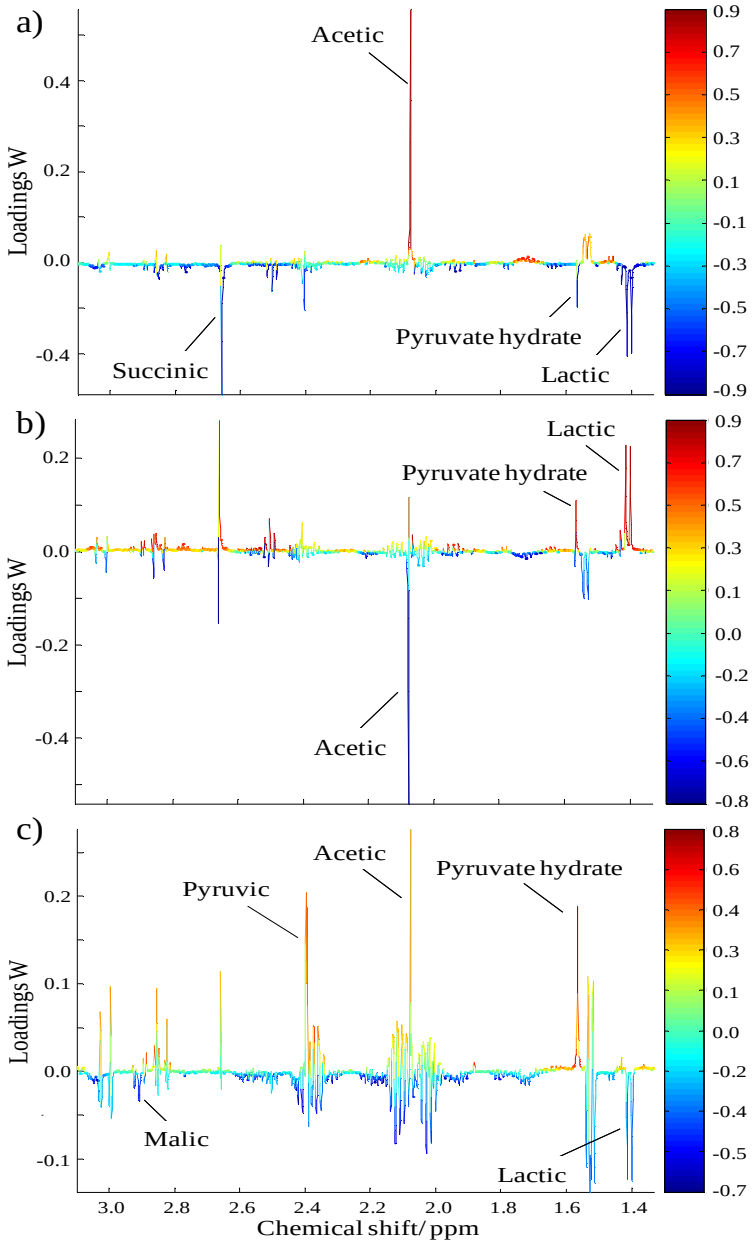

Figure 4.5. Loadings weights (**w1**) (LV1) plotted as a function of correlation for a) acetic acid and b) lactic acid, by PLS1-NMR/ CE direct, and c) pyruvic acid, by PLS1-NMR/ CE indirect. Color plot reflects the correlation between each NMR signal and the organic acid content: red and blue for extreme positive and negative correlations, respectively.

To assess the applicability of the developed PLS1-NMR regression models compared to those of the reference methods (CE/indirect detection and enzymatic) applied directly, the organic acid concentrations determined by each method were compared for one same beer set, test set D (Table 4.1, Figure 4.6).

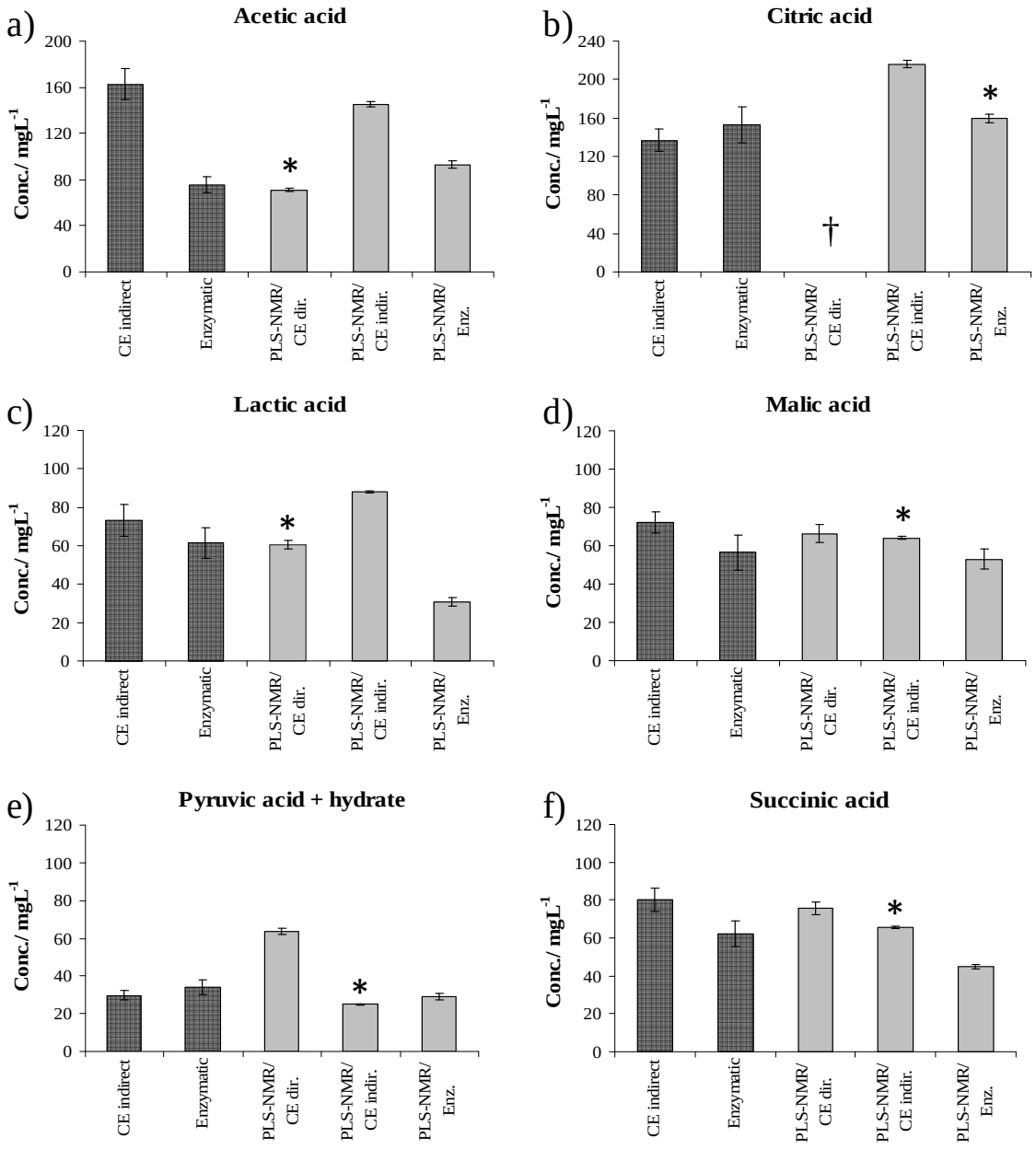

Figure 4.6. Column plots of organic acid concentrations obtained by the direct application of the reference methods CE/indirect detection and enzymatic tests (darker bars) and the PLS1-NMR models developed (PLS1-NMR/CE direct, PLS1-NMR/CE indirect and PLS1-NMR/Enzymatic). †: model not developed due to difficulties in citric acid detection by CE indirect detection. *: PLS1-NMR model showing best performance with basis on the parameters shown in Table 4.3.

Considering the best performance PLS1-NMR regression models (noted with * in Figure 4.6), as defined according to the model parameters in Table 4.3, it is clear that the PLS1-NMR results approach those obtained by the reference methods (CE/indirect and enzymatic essay), except for acetic acid for which a large positive deviation is noted for the results obtained with CE/indirect detection. This is probably due to the peak broadening and overlap problems found in the quantification of this acid by this analytical method, as noted before (Figure 4.1). Indeed, observing the column plots corresponding to the PLS1-NMR (Figure 4.6, grey bars) it is noted that the best performance models are the ones showing similar results when compared with the reference models bars (Figure 4.7, darkers bars).

Comparing the quantification results obtained for each organic by both the reference methods (CE with indirect detection and enzymatic essays) as well as by the best PLS1 regression model with the literature values, one can see that all organic acids concentrations are within the expected content range. Furthermore, to assess the quantification ability of CE with indirect detection and enzymatic essays, recovery tests were performed showing values of recovery between 75-114% for enzymatic essays and 84-118% for CE with indirect detection (except for acetic acid, for which the recovery tests were not perform). This slightly poorer recovery ability (> 20% variation) seen for enzymatic essays method reflects its already mentioned reproducibility problems.

The comparison of the quantification ability between PLS1-NMR regression models and NMR integration methods was carried out on a second test set of beer samples (test set E, Table 4.1). In this study, quantification by NMR integration vs. TSP and by ERETIC method were performed, as well as by the PLS1-NMR regression method selected with basis on its performance (method marked with * in Figure 4.6). Comparing the results obtained by integration vs. TSP and by ERETIC (Figure 4.7), it becomes clear that the agreement is satisfactory, with a concentration variation between both methods < 10 %, except for malic and pyruvic acids (Figure 4.7d,e) for which ERETIC gives relatively higher concentrations. In the case of pyruvic acid, this is justifiable by an underlying resonance from proline, possibly made more significant at the higher field at which ERETIC measurements were carried out (at 600 MHz spectrometer). It is possible that a similar effect occurs for the less intense resonances arising from malic acid.

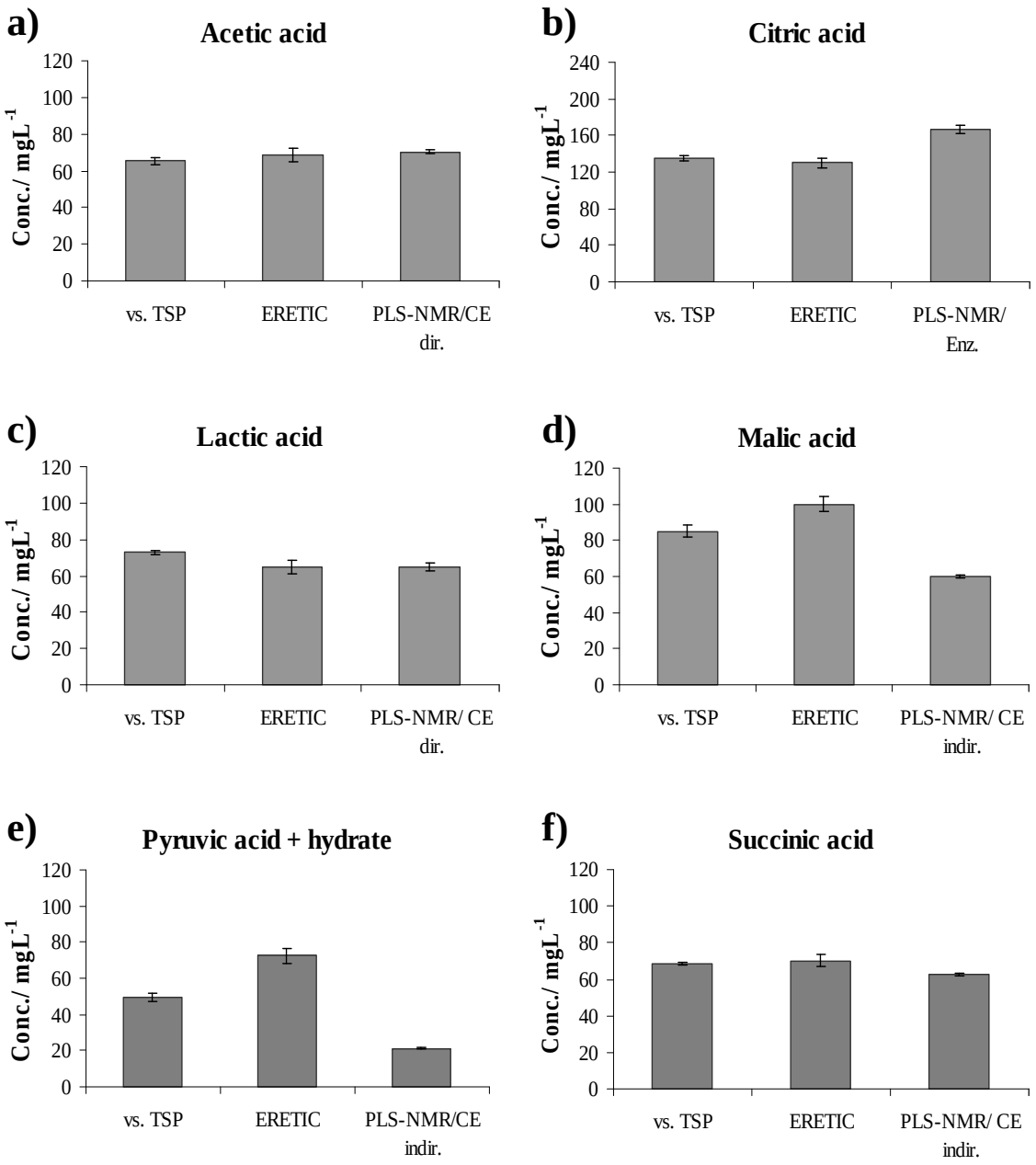

Figure 4.7. Column plots of organic acid concentrations obtained by NMR integration vs. TSP, the ERETIC method and the PLS1-NMR model showing best performance (* in Figure 4.6).

Comparison of the results obtained through peak integration with those obtained by the selected best performance PLS1-NMR regression models show that good agreement exists for acetic, lactic and succinic acids (Figure 4.8). Again, in the case of malic and pyruvic acids (Figure 4.8d,e), the apparent overestimation given by the integrals is consistent with the suggestion advanced above that underlying resonances affect the resonances of these acids. In the case of citric acid, the reason for the

discrepancy between integration and the PLS1 method is less clear and an overestimation by PLS1-NMR may not be ruled out. Furthermore, recovery tests were performed for the method of integration vs. TSP and the results show good recovery values, between 105-119%.

4.5 Conclusions

In this chapter, different NMR methodologies have been used to quantify the six main organic acids found in beer (acetic, citric, lactic, malic, pyruvic and succinic) and their performance compared. PLS1-NMR regression models were built using different reference analytical methods: CE, both with direct and indirect UV detection, and enzymatic essays. The regression models found to perform better were those obtained using CE indirect detection and enzymatic essays and their predictive power was compared with the results obtained through NMR integration methods, either using an internal reference or the ERETIC method. Results show that NMR integration methods are in good agreement with the PLS1-NMR models, except for malic and pyruvic, for which integration overestimates concentrations (probably due to additional underlying resonances), and for citric acid, for which an apparent overestimation by PLS1-NMR is observed. Indeed, by comparing the two integration methods (vs. TSP and by ERETIC), results also show a good agreement except for the same two compounds, where an overestimation of malic and pyruvic acid was seen for ERETIC method. These variations occurring on the integration methods indicate the strong influence of overlap in those signals integration. Furthermore, the concentration values obtained for all the developed quantitative methods show to be in agreement in the expected range of concentration found in literature for the studied acids.

Overall, the less time consuming method of PLS1-NMR is found to be suitable for organic acid quantification in beer, as long as the best combination of organic acid and particular PLS1-NMR model (differing in the reference method used for NMR calibration) is employed: PLS1-NMR/CE direct for acetic and lactic acids, PLS1-NMR/CE indirect for malic, pyruvic and succinic acids and PLS-NMR/enzymatic for citric acid.

5. STUDY OF THE CHEMICAL CHANGES ACCOMPANYING BEER FORCED AGING

5.1 Introduction

The organoleptic stability of beer during storage has been a major concern in the brewing industry (Guido *et al.*, 2003; Vanderhaegen *et al.*, 2006; De Schutter *et al.*, 2008). In fact, consumer acceptability of a specific type of beer depends on a complex network of chemical reactions occurring during storage, which may result in significant flavor degradation. The rates at which these reactions occur are affected by both intrinsic factors (e.g. oxygen content, pH, antioxidants contents and precursors concentrations of key odorants) as well as extrinsic factors (such as packaging, temperature and light). In order to study the impact of these parameters, forced aging protocols have been largely used, with thermal deterioration being widely employed to mimic the natural aging process (Chevance *et al.*, 2002; Varmuza *et al.*, 2002; Guido *et al.*, 2004; François *et al.*, 2006).

In this thesis, the forced aging process by thermal treatment (at 45 ^{o}C for up to 18 days) was followed by several analytical techniques: NMR spectroscopy, GC-MS and MIR, all of them combined with MVA to assess the major variations occurring in beer during the aging process. The use of the different analytical techniques of NMR spectroscopy and GC-MS, working at different quantitative domains, was based on their great ability to provide compound structural information, enabling new insights into beer aging chemistry to be achieved and possible new compounds and potential markers of the process to be unveiled. MIR spectroscopy was used in an attempt to achieve a fast method for assessing the degree of beer aging, with future potential application in the industry (less expensive and ease to use technique, with great implementation at industrial level). Furthermore, the aging process was also followed by sensory analysis, with the objective to establish a bridge between the organoleptic changes occurring in beer during aging and the compositional changes identified by the analytical techniques. In the following sections, the results obtained by the several methodologies applied are discussed and compared: 5.3) NMR/MVA; 5.4) GC-MS/MVA; 5.5) MIR/MVA; and 5.6) data correlation using MVA, for NMR/GC-MS, sensory/NMR and sensory/GC-MS data.

5.2 Experimental section

5.2.1 Optimization of the experimental design for beer forced aging

Taking the literature into account, the usually followed forced beer aging protocols have considered up to 3 to 7 days of thermal treatment, at temperatures at 37 ºC to 45ºC (Gijs *et al*., 2002; Varmuza *et al*., 2002; Guido *et al*., 2004; François *et al*., 2006). Accordingly, in this work, studies were performed in order to select the best combination of experimental parameters to use in the final, extended analysis. In each study, a rapid overall analysis of the correspondent NMR data was performed.

In a first attempt, beer bottles were stored at 37 ºC (± 4 ºC) during 7 days, in a dark room. Compositional changes were detected between fresh and aged beers, however, low reproducibility between samples from the same day of aging was noted, possibly reflecting the high temperature variability in the room. To overcome this, a following study consisted in storing beers bottles at 37 (±1) ºC during 7 days, in an oven, in the dark. Although temperature parameter was well-controlled in this case, no significant compositional changes were identified here, suggesting that the changes occurring due to the forced aging were minimal or in less extent than bottles variability. To assess the effects of bottles variability, a set of beers was mixed and transferred to separated vials in order to homogenize the starting composition of beer. The rapid analysis by NMR data showed compositional changes occurring between beers from different aging degrees, but, once again, some variability associated with beers from the same day was observed (although in less extent than the ones identified in the first study). Since some variability was still noted, it was decided to use individual beer bottles (in order to facilitate the aging procedure when large numbers of beer samples are used and to have well-defined samples for sensory analysis) and increase the temperature of aging. Accordingly, beers were stored at 40 (±1) ºC during 5 days, in an oven, in the dark. The results obtained showed meaningful compositional differences between fresh and aged and an improvement in terms of variability within days of aging, therefore this set of parameters was chosen for the extended aging study.

In this way, a forced aging study was performed, consisting in the storage of beer bottles during 15 days, at a temperature of 40 (±1) ºC, in the dark. Although this extended period of time may not reflect the relevant practical storage conditions of beer, it was intentionally used here in order to enhance compositional changes, making them

more easily measurable and interpretable by the analytical techniques. In the future, the subsequent fine-tuning of the methodology used here into the early stages of aging is intended. The aging process was followed employing both analytical techniques of NMR, MIR and GC-MS, as well as sensory analysis. However, in that study, minimal compositional changes to be measurable in a reproducible way or in less extent than bottles variability were identified, which led to no significant aging trend being observed either by NMR/MVA and MIR/MVA. To overcome this, it was decided to reproduce the forced aging process using stronger thermal treatment (at 45 (±1) ºC) and an extended period of 18 days. The studies performed using those forced aging conditions (at 45ºC during 18 days) are discussed in the following section.

5.2.2 Beer samples and sample preparation

All beer samples were from the same brand (lager beer), produced on the same site and date, and were kindly donated by UNICER, Bebidas de Portugal. To induce beer forced aging, twenty-seven lager beer samples (330 mL bottles with crown cork sealing caps) were stored in an oven at a temperature of 45 (±1) ºC, in the dark, during 18 consecutive days. Three beer bottles per day were removed for analysis on days 0, 1, 2, 3, 5, 7, 10, 13 and 18.

NMR spectroscopy

Beer samples (10 mL) were degassed in an ultrasonic bath for 10 minutes and prepared to contain 10% of D_2O and 0.025% 3-(trimethylsilyl)propionic-2,2,3,3-*d4* acid, sodium salt (TSP-*d4*) as chemical shift and intensity reference. Sample pH was adjusted to 1.90 ± 0.03 by adding 8-10 µL HCl 5% in D_2O, prior to sample transfer into 5 mm NMR tubes.

GC-MS

Beer samples were degassed for 10 minutes in an ultrasonic bath. Prior to GC-MS analysis, a beer extraction was performed by solid phase extraction (SPE) technique and using previously reported conditions for wine analysis (Ferreira *et al.*, 2003). After optimization of the method, the following conditions were employed: LiChrolut EN (500 mg) cartridges were conditioned by passing through 10 mL of pentane-

dichloromethane (20:1), 5 mL of methanol and 5 mL of an aqueous solution containing 5% (v/v) of ethanol. 50 mL of beer sample, containing 50 µL of internal standard 3-octanol (concentration of 460 mg L^{-1}), was passed through the SPE cartridge bed at a speed lower than 2 mL min^{-1}. The analytes were then eluted with 6 mL of dichloromethane. The extract was dried by adding anhydrous sodium sulphate. All chemicals employed were of analytical grade: dichloromethane and sodium sulphate anhydrous (Panreac), ethanol and methanol (Merck), n-pentane (Lab-Scan), 3-octanol (Sigma-Aldrich).

To assess the applicability of the SPE sample preparation to a non-wine matrix, e.g. beer, the following parameters were studied: repeatability, reproducibility and linearity of known aging markers. Repeatability (based in three consecutive SPE extractions) and reproducibility (consisting in extractions of the same beer sample in five different days) studies showed compounds variations percentages lower than 10%. Furthermore, linearity studies were also performed for known beer aging markers, i.e. methional and phenylacetaldehyde, based on the addition of known concentrations of those compounds (range from 2-50 μgL^{-1}) in beer matrix, with the corresponding calibration curves showing r values > 0.98.

MIR spectroscopy

For MIR measurements, sample preparation consisted simply of beer degassing in an ultrasonic bath, for 10 minutes.

5.2.3 Analytical measurements

NMR analysis

NMR spectra were recorded on a Bruker Avance DRX 500 spectrometer equipped with an actively shielded gradient unit with a maximum gradient strength output of 53.5 G/ cm and a 5 mm inverse probe, at 300 K. For each sample, two ^{1}H NMR spectra were acquired using the noesypr1d pulse sequence (Bruker pulse program library) with water presaturation. 128 transients were collected into 32768 (32 K) data points, a spectral width of 8013 Hz, acquisition time of 2.0 s and relaxation delay of 5 s. Each FID was zero-filled to 64 k points and multiplied by a 0.3 Hz exponential line-broadening function prior to Fourier transform. Spectra were manually baseline corrected and

chemical shifts referenced internally to the TSP peak. All peak assignments were carried out as described in section 3.3.

Peak integration was performed using AMIX.3.9.5, Bruker BioSpin and their areas were normalized relatively to the total spectral area. For evaluation of dextrins changes, integration of specific anomeric signals was performed according to previous reports (Duarte *et al*, 2003), as described in section 3.3.

GC-MS analysis

For GC-MS analysis, a gas chromatograph, Varian 450 GC, equipped with a mass detector, Varian 240-MS, and an ion trap analyzer were used. The injection port was lined with 0.75 mm I.D. splitless glass liner working at 220 ºC. The split valve was opened during 0.5 min after injection. Volatiles were separated in a VF-WAX (15 m x 0.15 mm x 0.15 µm) capillary column from Varian. Two different oven temperatures were used:

i) Conventional runs (optimum chromatographic separation) with the oven temperature held at 40 ºC for 1 minute, and then increased at a rate of 5 ºC min^{-1} to a final temperature of 230 ºC, which was held for 2.5 minutes, with a total run time of 41.5 minutes.

ii) Rapid runs consisting in holding the oven temperature at 70 ºC for 1 minute and then increasing at a rate of 25 ºC min^{-1} to a final temperature of 240 ºC (held for 5 minutes), with a total run time of 12.8 minutes.

The mass spectrometer operated in the electron impact (E.I) mode at 70 eV, scanning the range *m/z* 33 – 350. The electron ionization parameters, ionization time of 200 µsec and a maximum ion time of 5000 µsec, were fixed from 6.0 to 41.5 minutes and from 2.0 to 12.8 minutes for conventional and rapid runs, respectively, to have the same number of scans in each analyzed sample.

Peaks integration was carried out with MS Data Review, Varian MS workstation and their areas were normalized relatively to the area of the internal standard area (3-octanol).

MIR analysis

All MIR spectra were collected on a FT-IR Perkin Elmer BX spectrometer, using a single reflectance horizontal ATR cell (Golden Gate, equipped with a diamond crystal),

in a controlled temperature (23 (± 3) ºC) and humidity (35 (± 8) %) room, after optimization. Two methodologies were carried out for beer analysis, termed direct and droplet evaporation methodologies:

i) Direct spectra: 10 μL of beer sample was placed in the ATR crystal and each spectrum was recorded in the spectral range of 4000-600 cm^{-1}, by accumulating 64 scans with a resolution of 4 cm^{-1}. For each sample, a total of five spectra were acquired.

ii) Droplet evaporation spectra: 6 μL of beer sample was used (corresponding to the minimum volume of sample in order to cover completely the ATR crystal) and each spectrum was recorded with the same parameters as the ones used in the direct MIR. The time needed for the droplet evaporation was approximately 20 minutes, after which five replicas, consisting in five consecutive measurements of the dried sample, were acquired. Both repeatability (three consecutive beer droplet evaporations and consequent MIR analysis) and reproducibility (droplet evaporation and subsequent analysis in three different days) of the droplet evaporation methodology were assessed by analyzing three different beer samples. By visual analysis of beer spectra as well as by combining MIR spectra with PCA, both reproducibility and repeatability, the latter in less extent, show some degree of variation, indicating variability associated to the evaporation process.

Between each sample measurement, the ATR crystal was cleaned with water and the background spectrum was acquired using the same instrumental conditions than those described above and used for sample analysis. To avoid possible intensity fluctuations due to minor variations of room temperature and humidity, the samples corresponding to each studied day of aging were analyzed randomly. All spectra were converted into .DX format for statistical analysis.

Sensory analysis

For sensory exploratory analysis, four tasters of the internal sensory panel of UNICER, Bebidas de Portugal were recruited based on their good sensory ability and experience on the identification of stale off-flavors. The sensory analysis took place on the morning at an adequately isolated room. Panelists were asked to comment on the general quality of beer (discrimination test). All beers were tasted at 10 ºC, in random order by blind analysis (panelists and organizers of the sensory analysis had no previous access to beer codes). The degree of beer staling was evaluated using a seven-point

sensory score scale ranked from 0 (no sign of staling) to -3 (very strong level of staling), sub-divided by a 0.5 scale and with a threshold limit of -1.5 level for rejecting the beer.

5.2.4 Multivariate analysis (MVA)

Application to NMR data

For multivariate analysis, each spectrum was divided into variable width "buckets", using Amix-Viewer 3.9.2 software. The water (4.7-4.9 ppm) and ethanol (1.0-1.3 ppm and 3.6-3.7 ppm) resonance regions were excluded from the matrix, as well as the 0.0-0.8 ppm region where no sample signals resonate. PCA and PLS-DA models were performed considering either the total spectral window (0.8-10.0 ppm) or the aliphatic (0.8-3.1 ppm), anomeric (4.5-5.8 ppm) and aromatic (5.8-10.0 ppm) regions separately. Each spectrum was normalized by dividing each bucket by the total spectral area and each variable (bucket) was i) autoscaled (i.e. unit variance) or ii) column centering, using SIMCA-P 11.5 software. For interpretation purposes, the loadings profiles obtained (when using the autoscaled procedure) were back-transformed by multiplying all variables (buckets) by their standard deviation and their plots were color-coded as a function of the variable importance on the projection (VIP), being obtained using R-statistical software 2.11.0 and Plotrix package (Lemon, 2010).

2D correlation analysis of the NMR data was performed on either whole spectra or spectral regions (aliphatic and aromatic). Each spectrum (or sub-region) was normalized by total spectral (or sub-region) area and the average spectra corresponding to each aging day were computed. 2D correlation analysis was performed using the spectrum of fresh beer (day 0) as reference and all models were built using an in-house application co-developed by the University of Aveiro and the AgroParis Tech, France, the synchronous and asynchronous contour maps having been produced with MATLAB 7.8.0. These maps provide graphical representations of the quantitative relationships between signal intensities for all pairs of spectral variables (NMR data), as a function of a given perturbation (Kirwan *et al.*, 2008; Kirwan *et al.*, 2009), in this case, beer aging. The correlation between signals is treated as a complex number comprising two orthogonal components (synchronous, Φ, and asynchronous, Ψ, correlations matrices), which provide complementary information (Noda and Ozaki, 2004). The synchronous matrix represents simultaneous or coincidental changes, expressed as positive (in the

same direction) or negative (in the opposite direction) correlations between spectral data points, whereas the asynchronous matrix represents sequential (or *"out of phase"*) changes in spectral intensities. Briefly, for pairs of variables (data points) positively correlated in the synchronous matrix, the sign of an asynchronous peak is positive if the intensity change at a variable occurs before (or to a higher extent) that of a second variable in the sequential order. On the other hand, the sign of the asynchronous peak becomes negative if the change at the first variable occurs later (or to a lesser extent). This sign rule is reversed if the synchronous correlation intensity at the same coordinate is negative. In addition, pairs of variables identified in the synchronous matrix alone (i.e. producing no signal in the asynchronous map) are said to be correlated due to following the *same* variation rate during the process under study (Noda and Ozaki, 2004).

Application to GC-MS data

Two types of GC-MS data were acquired and processed for multivariate analysis:

i) Conventional runs: the chromatographic domain used was set between 7.0 and 33.5 minutes (1332 scans). The range of *m/z* values used ranging from 39 to 250, as no significant information was detected off that *m/z* range. The *m/z* values related to chemical noise characteristic of column and solvent, i.e. 73 and 221 *m/z*, were excluded giving a total of 278388 (1332 x 209) points per analysis. For each day of analysis, 3 beer samples were analyzed, without replicates, giving a total of 27 GC-MS analyses. Therefore, the size of GC-MS data matrix was 7516476 (278388 x 27) points.

ii) rapid runs: chromatographic domain ranging from 2.25 to 8.00 minutes (275 scans). The *m/z* range used was the same as above, excluding *m/z* values 40, 73 and 221 (chemical noise characteristic of column and solvent), giving a total of 57200 (275 x 208) points per analysis. As two replicates were analyzed per sample, a total of 54 analyses were performed and the matrix data size for rapid GC-MS had 3088800 points (57200 x 54). Although the rapid GC-MS data had two replicates for each sample, the corresponding matrix had less than half the number of points of the conventional one.

Each GC-MS sample data is inherently a matrix, with one dimension related to the chromatographic information (retention time), and the other related to the mass spectra (*m/z* pattern fragmentation). Therefore, each sample consisted of $\mathbf{S(c)_i}_{(1332x209)}$, for conventional runs, and $\mathbf{S(r)_i}_{(275x208)}$, for rapid runs, matrices, where *i* corresponds to the

number of samples (27 for conventional and 54 for rapid runs). For multivariate analysis, each matrix $\mathbf{S_i}$ was unfolded to give a vector $\mathbf{s}^{T}_{\mathbf{i}\ (1,\ j)}$, where *j* corresponds to the number of variables of each analysis. Then, all vectors were row concatenated to give a $\mathbf{Q}_{(i,\ j)}$ matrix. This **Q** matrix was then decomposed by PCA into:

$$\mathbf{Q}_{(i,\ j)} = \mathbf{T}_{(i,\ k)} \,.\, \mathbf{P}^{T}_{(k,\ j)} \,.\, \mathbf{E}_{(i,\ j)} \qquad \text{[Eq. 5.1]}$$

where **T** and **P** represent the scores and loadings matrices, respectively, **E** holds residual unexplained variation and *k* corresponds to the extracted number of principal components. To facilitate interpretation, each one of loadings columns *k* (a vector) were folded back to give a matrix, which was depicted as a 2D map:

$$\mathbf{p}_{k(j,\ 1)} \rightarrow \mathbf{P}_{k(l,\ m)}, \text{ where } j = l \,.\, m \qquad \text{[Eq. 5.2]}$$

such as *l* and *m* are, respectively, the number of scans (expressed as retention time) and the *m/z* values correspondent to each sample. PCA was performed using MATLAB 7.8.0 and an in-house application co-developed by the University of Aveiro and the AgroParisTech, France.

Application to MIR data

Two types of MIR data were acquired and processed for chemometric studies:

i) Direct spectra: only the region between 1200-850 cm^{-1} (*fingerprint region*) was studied, avoiding the major influence of the water absorption bands, located at 1800-1500 $cm^{-1}$and 3800-3000 cm^{-1}.

ii) Droplet evaporation spectra: two spectral regions were studied, namely the 3000-850 cm^{-1} (excluding the carbon dioxide (CO_2), between 2500-2200 cm^{-1}, and the water, located at 1800-1500 cm^{-1}, regions) and the 1500-850 cm^{-1} regions.

The average spectra corresponding to each beer sample were computed for both types of obtained MIR data and two data processing approaches were followed: i) spectra autoscaled (mean centered and variance scaling); and ii) application of a 2nd derivative transformation, using the Savitzky-Golay procedure (Savitzky and Golay, 1964), and, then spectra autoscaled. The Savitzky-Golay procedure was used to compute 2nd derivative and perform smoothing at the same time. The use of this derivative transformation can partially compensate for baseline offset between samples and reduce instrument drift effects (Wang *et al.*, 2009).

Both PCA and PLS-DA were performed using MATLAB 7.8.0, SIMCA-P 11.5 software and an in-house application co-developed by the University of Aveiro and the AgroParisTech, France.

Data fusion by outer-product analysis (OPA) and application of statistical heterospectroscopy (SHY) between NMR and GC-MS data

Principal component transform-outer product analysis in the context of partial least square-discriminant analysis (PCT-OPA-PLS-DA) was performed between NMR and GC-MS data, with the goal of determining the existing relations (both linear and non-linear relations) between those data sets.

Prior to the data fusion by OPA, PCT methodology was applied to each data set, NMR (866 points x 54 spectra) and GC-MS (57200 points x 54 runs), in order to decompose each data set into principal components (PC's) corresponding to the most important sources of variability (54 x 54). This data decomposition methodology was successfully applied in the past (Barros *et al.*, 2004; Barros *et al.*, 2007; Barros *et al.*, 2008), where it was demonstrated that if sufficient PC's are extracted (i.e. if the number of PC's is equal to the rank of the studied matrices) all sources of variability are preserved. Briefly, a matrix **X** can be decomposed into orthonormal (linear-independent) matrices based on the calculation of the vectors or axes, commonly called principal components (PCs), corresponding to the most important sources of variability present in matrix **X**:

$$\mathbf{X}_{(n,\,m)} = \mathbf{T}_{(n,\,k)}\,\mathbf{P}^{\mathbf{T}}_{(k,\,m)} \,.\, \mathbf{E}_{(n,\,m)} \qquad \text{[Eq. 5.3]}$$

where *n* is the number of objects (rows), m the number of variables (columns) and *k* the number of PCs. If sufficient PCs are extracted all sources of variability are preserved, condition that is fulfill if the number of PCs is equal to the rank of **X**, i.e., if $k = n$. Therefore, to compress the original NMR and rapid GC-MS data sets, full rank PCA models were built for both matrices (previously normalized by rows), i.e., 54 vectors (PCs) were calculated corresponding to the number of analyses made in each matrix. The correspondent scores matrices, $\mathbf{T}_{\text{GC-MS (54, 54)}}$ and $\mathbf{T}_{\text{NMR (54, 54)}}$, were recovered and the outer-product (OP) between those matrices was performed, giving a final matrix **H:**

$$(\mathbf{H}_{(54,\,2916)} = \mathbf{T}_{\text{GC-MS (54, 54)}} \otimes \mathbf{T}_{\text{NMR (54, 54)}}) \qquad \text{[Eq. 5.4]}$$

where $\otimes$ represents the outer-product operator. The resulting OP matrix **H** was mean-centered and PLS-DA was applied. To ease the interpretation of the loadings profiles

(**p**) corresponding to the OP matrix **H**, each one of relevant loadings columns *k* (a vector) were folded back to give a matrix, $\mathbf{p}_{k(2916,\ 1)} \rightarrow \mathbf{P}_{k(54,\ 54)}$. The recovery of the loadings in the original variable space (allowing the interpretation of the relations between NMR and rapid GC-MS data), was obtained in a straightforward manner, by pre- and post-multiplying the OP loadings by the loadings matrices corresponding to GC-MS and NMR data, $\mathbf{P}_{\text{GC-MS }(57200,\ 54)}$ and $\mathbf{P}_{\text{NMR }(866,\ 54)}$, calculated previously. For each extracted PC subscript as *a*:

$$\mathbf{P}_{a\ (866,\ 57200)} = \mathbf{P}_{\text{NMR }(866,\ 54)} * \mathbf{P}_{\text{OPA }(54,\ 54)} * \mathbf{P}_{\text{NMR }(54,\ 57200)} \qquad \text{[Eq. 5.5]}$$

meaning that, to each NMR variable (in a total of 866) corresponds a vector *l* (with 57200 points) related to the GC-MS data. As GC-MS data is inherently bi-dimensional, one must fold back the vector *l* to give a matrix, which was depicted as a 2D map: $l_{(57200,\ 1)} \rightarrow L_{(275\ x\ 208)}$, expressing the retention time and the *m/z* values. PCT-OPA-PLS-DA operations were performed using MATLAB 7.8.0 and an in-house application co-developed by the University of Aveiro and the AgroParisTech, France.

Statistical heterospectroscopy (SHY) between NMR and GC-MS data was also performed, the detection of highly correlated signals from both domains. SHY consists in the built up of a correlation matrix **C** according to:

$$\mathbf{C}_{(866,\ 57200)} = [1/(n\text{-}1)] * \mathbf{X}_{1\ (866,54)}^{\mathbf{T}} * \mathbf{X}_{2\ (57200,\ 54)} \qquad \text{[Eq. 5.6]}$$

where $\mathbf{X}_1$ and $\mathbf{X}_2$ refer to the autoscaled matrices of NMR and GC-MS respectively; and *n* is the number of analyses per data set. Therefore, **C** is a matrix of *866 x 57200*, where each value is a correlation coefficient between two variables of the matrix $\mathbf{X}_1$ and $\mathbf{X}_2$ (NMR and GC-MS data). For each NMR variable, a correspondent correlation vector *l* (with 57200 points), related to the GC-MS data, is obtained. Folding back the vector *l*, a 2D matrix is obtained expressing retention time and the *m/z* values of GC-MS, which was depicted as a 2D map: $l_{(57200,\ 1)} \rightarrow L_{(275\ x\ 208)}$, corresponding to the correlations between NMR variable and GC-MS data. Correlations between the two matrices and corresponding correlation plots were performed using MATLAB 7.8.0.

PLS1 regression of sensory/ NMR and sensory/ GC-MS data

PLS1 regression models were developed to assess the relationships between the sensory and the analytical response, as a function of days of aging. For PLS1 regression models, using as reference (the **y** data vector) the sensory response of the UNICER,

Bebidas de Portugal panel, the data matrices **X** employed corresponded, for NMR data, to the spectral regions of aliphatic (0.8-3.1 ppm), anomeric (4.5-6.0 ppm) and aromatic (6.0-10.0 ppm) regions, as well as, for rapid GC-MS data, to the chromatographic domain from 2.25 to 8.00 mins and the *m/z* range between 39 to 250, excluding *m/z* values 40, 73 and 221.

All PLS1 regression studies were performed using software co-developed by the University of Aveiro and the AgroParisTech, France. To improve the interpretation of the PLS1 regression models, the loadings weights profiles were colored as a function of the correlation between each variable and the corresponding sensory response data (as expressed in the **y** data vector), being obtained using R-statistical software 2.11.0 and Plotrix package (Lemon, 2010) for NMR data and MATLAB 7.8.0 for GC-MS data.

5.3 Use of NMR spectroscopy combined with multivariate analysis (NMR/MVA)

In this section, multivariate analysis, namely PCA and PLS-DA, of the NMR data enabled the assessment of the main profile changes accompanying beer aging, allowing specific relevant compounds to be identified and deeper knowledge of the process to be obtained. In addition, the multivariate analysis technique of 2D correlation analysis (Noda, 1993; Noda and Ozaki, 2004) was applied to the NMR data, enabling the relevant compounds variations to be confirmed and inter-compound correlations to be assessed, some reflecting common metabolic/chemical pathways and, therefore, offering improved insight into the chemical aspects of beer aging.

5.3.1 ^{1}H NMR characterization of fresh and aged beer

Visual comparison of the spectra of fresh beer and that of beer aged for 18 days (Figure 5.1a,b) reveals a general similarity between both spectra profile, although some changes are encountered and highlighted by oval shapes. Insets of these major changes are illustrated in Figure 5.1c, revealing the clear increase in 5-HMF peaks (mainly visible at 6.67 and 9.45 ppm) and the decrease in the unassigned peaks at 1.46 and 6.35 ppm (named as unassigned 1 and 7, respectively). Indeed, 5-HMF is a well known product of sugar degradation, being described as increasing with time at an approximately linear rate and varies logarithmically with the storage temperature (Madigan *et al.*, 1998; Vanderhaegen *et al.*, 2006).

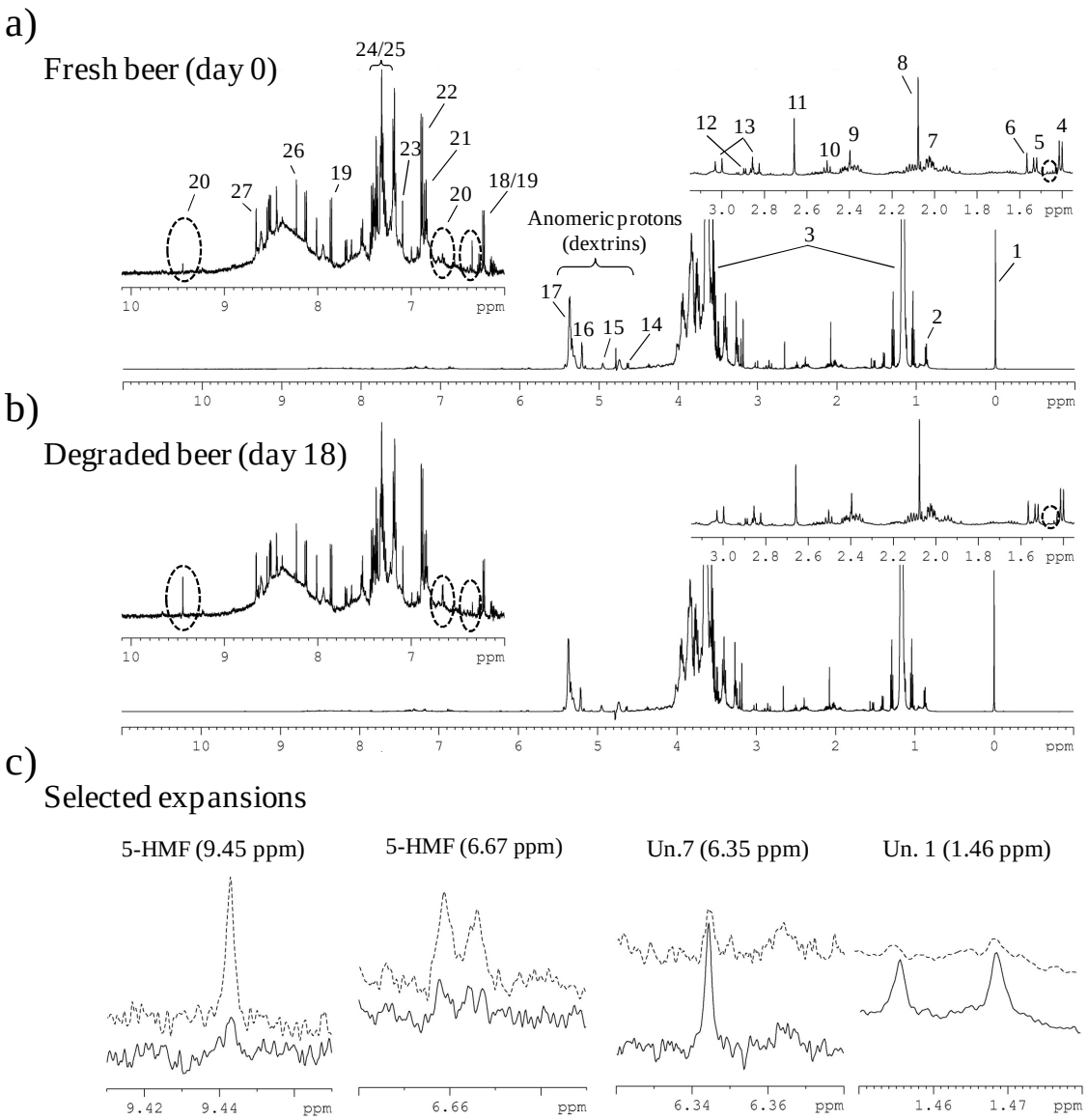

Figure 5.1. 500 MHz ^{1}H NMR average spectra of a) fresh beer (day 0) and b) beer degraded (18 days at 45 ºC), with insets shown for aliphatic (0-3 ppm) and aromatic (6-10 ppm) regions. Peak assignments: 1, TSP; 2, higher alcohols (isobutanol, isopentanol, propanol); 3, ethanol; 4, lactic acid; 5, alanine; 6, pyruvic acid, hydrate form; 7, proline; 8, acetic acid; 9, pyruvic acid; 10, γ-aminobutyric acid (GABA); 11, succinic acid; 12, malic acid; 13, citric acid; 14 – 17, anomeric protons (H1): β red., α (1→6), α red., α (1→4), respectively; 18, cytidine; 19, uridine; 20, 5-HMF; 21, tyrosol; 22, tyrosine; 23, gallic acid; 24, phenylalanine; 25, phenylethanol; 26, formic acid; 27, histidine. Major signal variations between fresh and aged beer are highlighted by oval shapes and shown in c): un. (unassigned) 1 (1.46 ppm), un. 7 (6.35 ppm) and 5-HMF (6.67 and 9.45 ppm); dashed lines: aged beer (day 18 at 45 ºC); and solid lines: fresh beer (day 0).

2D NMR experiments were performed for a more detailed characterization of the beer matrix correspondent to fresh beer (day 0) and degraded beer (day 18). Although a different overall profile between them was observed, no additional information about the unassigned signals was obtained in these experiments, when compared to the description of beer chemical composition described in Chapter 3.

5.3.2 Multivariate analysis of NMR data (NMR/MVA) as a function of aging

To investigate the main sources of variability in the ^{1}H NMR spectra corresponding to the accompanying of beer aging process, PCA and PLS-DA models were built for the whole spectra (0.8-10.0 ppm), excluding the water and ethanol resonances. Figure 5.2a shows the 2D and 3D PCA scores scatter plots obtained for the NMR data. The typical 2D plot PC1 vs. PC2 (Figure 5.2a, left), accounting for, respectively, 20% and 14% of the total variability, shows no clear separation between beer differing in aging extent, with the exception of samples corresponding to days 13 and 18, which are shifted to negative PC2. However, the 3D representation (PC1 vs. PC2 vs. PC3, which together account for 42 % of the total variability) (Figure 5.2a right) shows clear sample separation, revealing three main sample groups: days 0 to 5, days 7 and 10 and days 13 and 18.

The corresponding PLS-DA models confirms this separation in both 2D and 3D representations (Figure 5.2b), showing improved separation of day 5 samples within the earlier aging stages, and reveal four distinct sample groups (days 0-3, day 5, days 7 and 10, and days 13 and 18). This PLS-DA model is characterized by R^2X and R^2Y (expressing the explained variance of X and Y, respectively) values of 0.49 and 0.42, respectively, and a Q^2 (expressing the predictive ability of the model) of 0.23, when using five latent variables. By performing a permutation test, the robustness of this model was confirmed. Comparing both MVA methods (i.e. PCA and PLS-DA), one has decided to use the PLS-DA approach in this study of beer aging, as it showed greater discriminative potential in terms of days of aging. Examination of each PLS-DA scores plot, unveiled a clear aging trend in the main latent variable, [**t1**], with the last days of aging (days 13 and 18) being identified as separated and located in negative [**t1**]. Moreover, in [**t2**] axis, a separation between days 7 and 10 of aging from the rest of the studied days was also noted, with the former being identified as located in negative [**t2**].

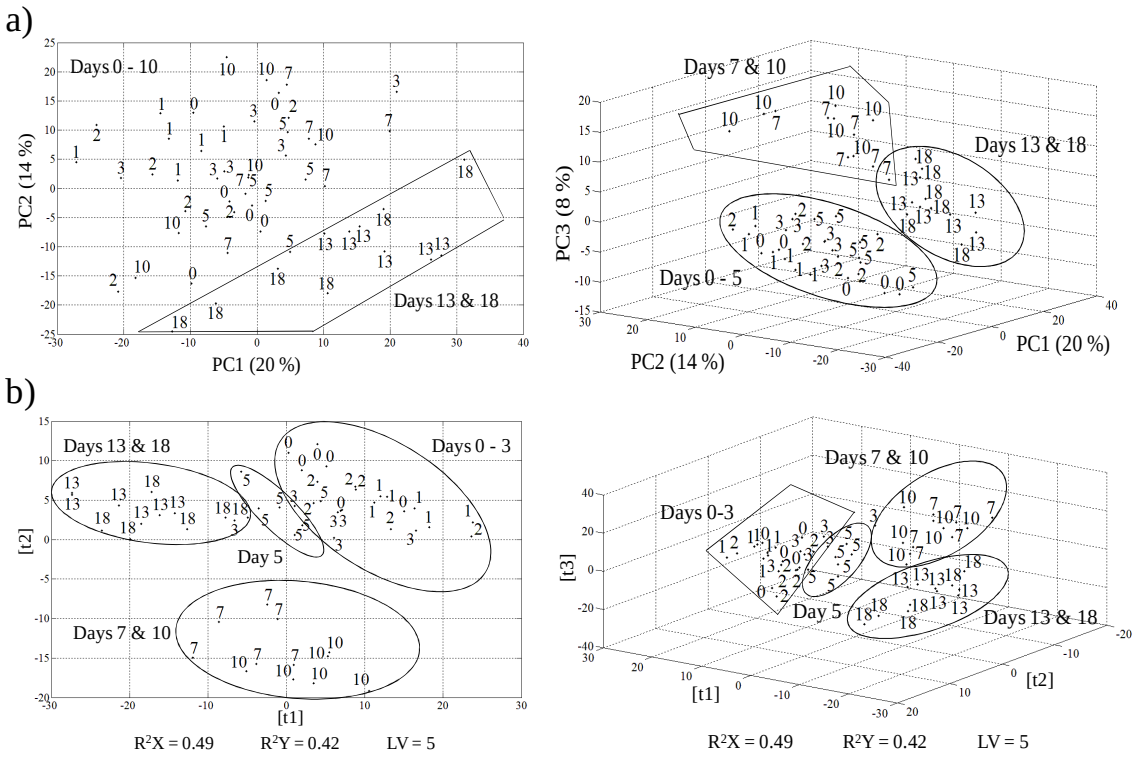

Figure 5.2. a) PCA and b) PLS-DA scores scatter 2D and 3D plots obtained for whole NMR spectra as a function of thermal forced aging. Numbers indicate aging days and grouping shapes were drawn to help the identification of membership.

Figure 5.3a shows the loadings weight (**w1**) profile corresponding to the main latent variable [**t1**], where the colored scale expresses the variable importance on the projection (VIP) of each data point, with the more relevant variables being shown in red. In this plot, negative signals reflect metabolites showing concentration increases with aging (mainly on days 13 and 18), whereas positive peaks reflect contents decreasing with aging. The loadings profile suggests overall increases in 5-HMF, acetic, pyruvic and succinic acids, branched dextrins (viewed by H1 resonances at signals 13 and 15, respectively for α (1→6) and α (1→4) in branched oligomers) and unassigned 2, 3, 5 and 6. In addition, decreases are suggested for citric acid, GABA, gallic acid, linear dextrins (viewed by H1 resonances at signals 12, 14 and 16 for β red., α red. and α (1→4), respectively, in linear oligomers) and unassigned signals 1, 4, 7, 8 and 9. Furthermore, Figure 5.3b shows the loadings weight (**w2**) plot, where the negative peaks are related with signals more abundant on days 7 and 10, suggesting that the beers of days 7 and 10 have higher amounts of acetic, pyruvic and its hydrated form and succinic acids as well as a decrease on the ratio linear/branched dextrins.

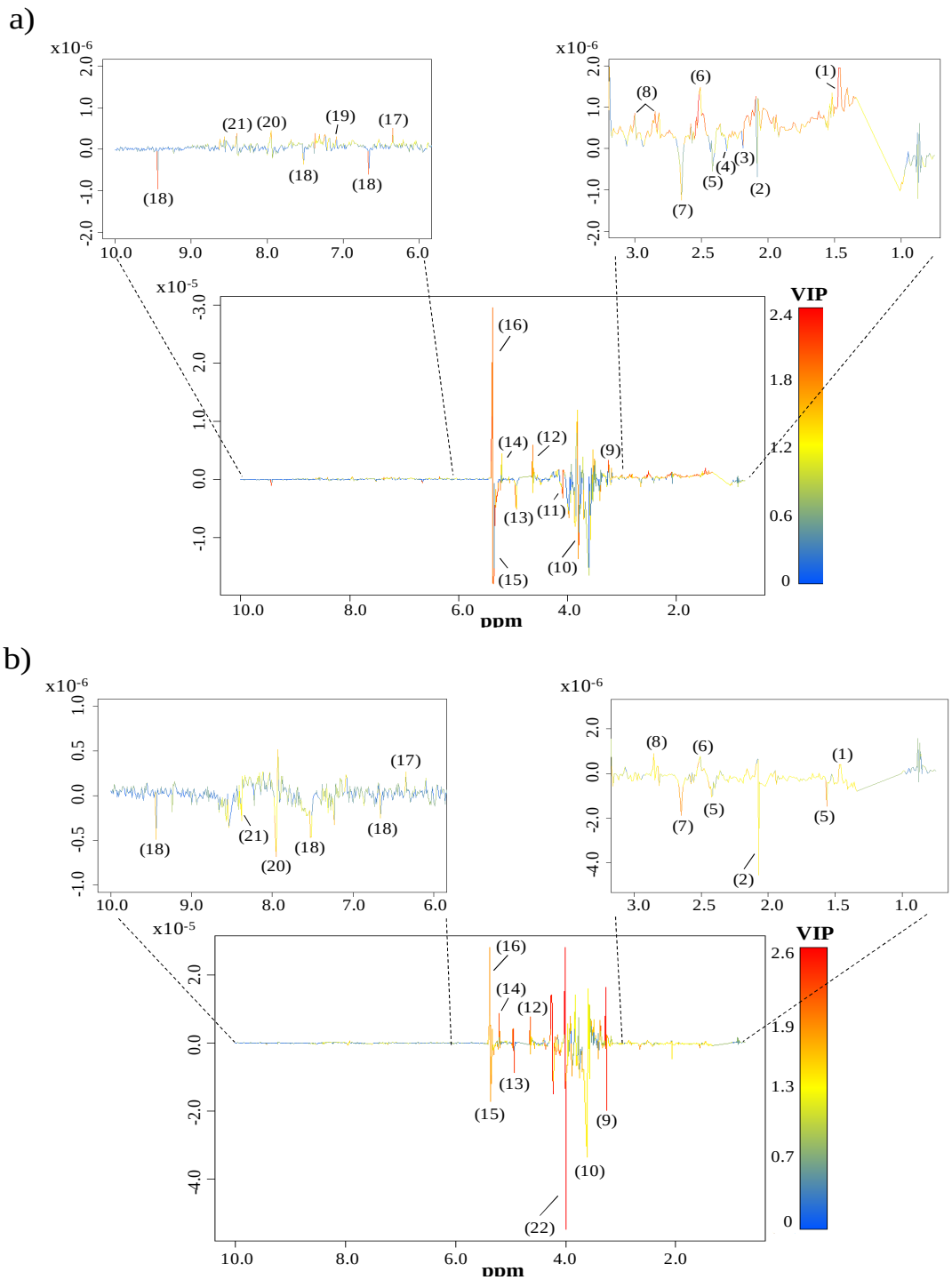

Figure 5.3. PLS-DA loadings weight plot corresponding to a) (**w1**) and b) (**w2**) obtained for whole NMR spectra. Signals varying with aging refer to (1) un. (unassigned) 1, (2) acetic acid, (3) un. 2, (4) un. 3, (5) pyruvic acid and hydrate form, (6) GABA, (7) succinic acid, (8) citric acid, (9) un. 4, (10) un. 5, (11) un. 6, (12) H1β (red.), (13) H1α (1→6), (14) H1α (red.), (15) branched H1α (1→4), (16) linear H1α (1→4), (17) un. 7, (18) 5-HMF, (19) gallic acid, (20) un. 8, (21) un. 9 and (22) un. 10. The color scale reflects the variable importance (VIP) of each data point.

In order to perform a more detailed analysis, PLS-DA models of the aliphatic, sugar and aromatic spectral regions were built and analyzed separately. The corresponding PLS-DA scores plots (Figure 5.4, left side) indicate that changes in aliphatic compounds are noted as early as day 3 (Figure 5.4a), whereas day 5 beer shows compositional similarity to day 7 beer in terms of aliphatic components. Regarding sugar composition (Figure 5.4b), separation of day 3 beer from earlier aging stages is observed, with day 3 showing compositional similarity with day 5. A similar separation trend is also noted in terms of the compounds resonating in the aromatic region (Figure 5.4c). Therefore, the above results show that NMR/MVA of spectral regions may provide finer differentiation between earlier aging stages (namely, within days 0 and 3), information which may become useful for practical evaluation of beer aging.

The analysis of the corresponding loadings plots (Figure 5.4, right) enabled the confirmation of the main compositional variations tendencies mentioned in the NMR/PLS-DA results considering the whole spectra, namely 5-HMF, acetic, pyruvic and succinic acids and unassigned signals 2 and 3 increase and decreases in the ratio linear/branched dextrins, citric and gallic acids, GABA and unassigned signals 1, 7, 8 and 9. In addition, higher alcohols (0.86-0.89 ppm) and formic acid (8.23 ppm) were also suggested as increasing with the aging process. Assessment of PLS-DA results was then attempted by integration of the identified compounds so that variations could be either confirmed or discarded. In this way, 10 assigned compounds, dextrins and 5 still unassigned peaks (Table 5.1) were identified as varying significantly with thermal aging. Although citric acid and four unassigned compounds (unassigned 2, 4, 8 and 9) had been suggested by PLS-DA to vary, such variations were found not relevant, by signal integration, possibly due to low peak intensity (and hence relatively high uncertainty in the integration procedure) and/or apparently random variations during aging. Some examples of integral variations as a function of aging are shown in Figure 5.5 for a) 5-HMF, b) higher alcohols, c) GABA d) succinic acid and unknown compounds e) 1 and f) 7. The changes relating to dextrins are also shown, being expressed in terms of variations in the average number of branching points/molecule and the average number of glucose units/molecule (Figure 5.5g and 5.5h, respectively).

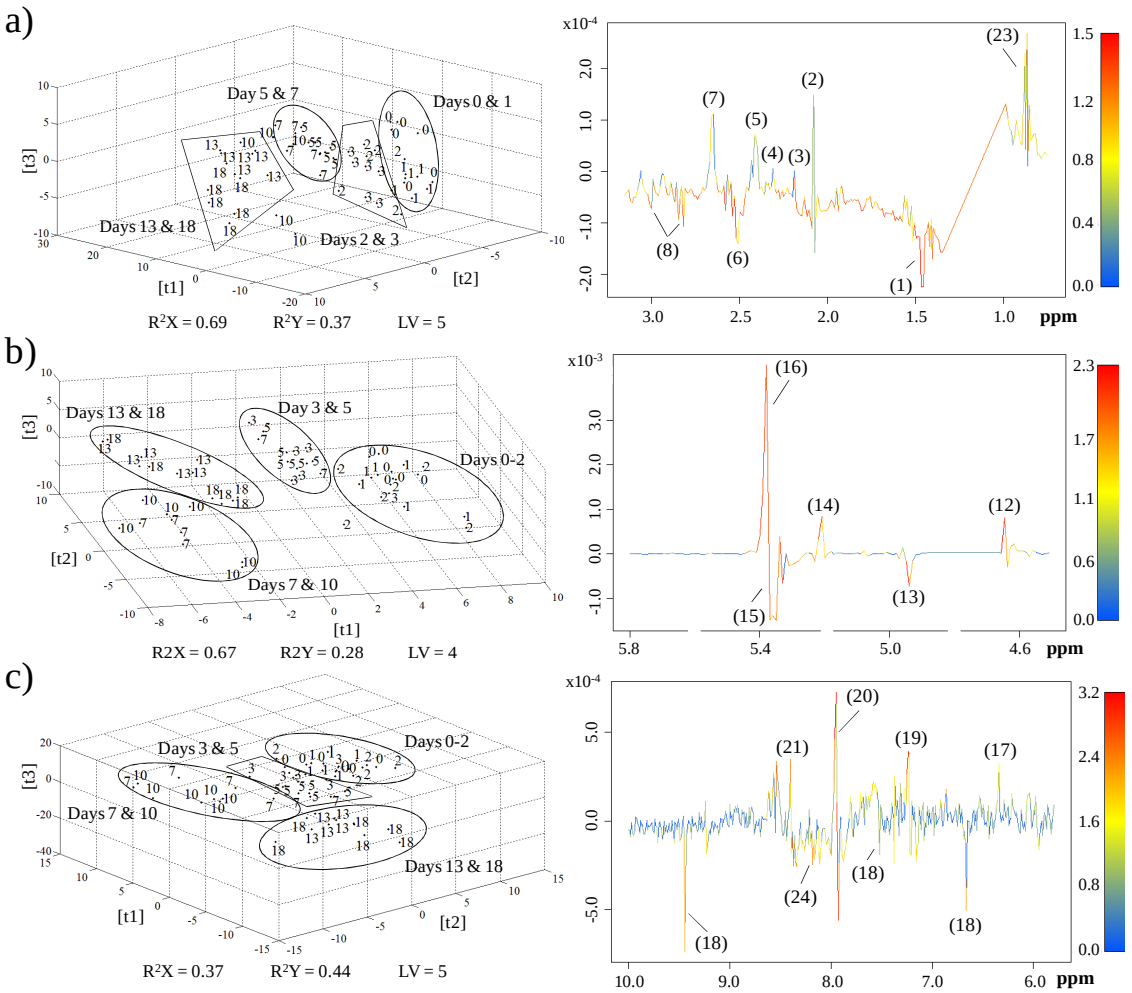

Figure 5.4. PLS-DA 3D scores scatter plots and corresponding loadings weight (**w1**) plot obtained for (a) aliphatic (0.8-3.1 ppm), (b) anomeric (4.5-5.8 ppm) and (c) aromatic regions (5.8-10.0 ppm), as a function of number of aging days (indicated by numbers shown). Grouping shapes were drawn in the scores scatter plot to help the identification of membership. Signals varying with aging are identified by the numbering shown in Figure 5.3, with the addition of (23) higher alcohols and (24) formic acid. The color scale in the loadings plot reflects the VIP of each data point.

The compositional changes measured by peak integration were consistent with the aging trend viewed through PLS-DA, with fresher beers (up to 3 days of aging) being distinguished by higher amounts of GABA, gallic acid, linear dextrins and unassigned 1 and 7, and lower amounts of 5-HMF, higher alcohols (weak tendency, not considering day 0), acetic, formic, pyruvic and succinic acids, branched dextrins and unassigned 3 and 6. Beers from day 5 onwards are then distinguished by a continuing increase in 5-HMF along with smaller increases in higher alcohols (weak intensity),

organic acids (acetic, pyruvic and succinic acids) and unassigned 3, 5 and 6, as well as decreases in GABA, unassigned 1 and 7 and the ratio linear/branched dextrins. These variations become gradually more marked as aging proceeds, although small decreases in acetic, pyruvic and succinic acids and unassigned 3 occur in the late stages of aging (days 13 and 18).

Table 5.1. List of metabolites or NMR resonances varying as a function of thermal aging and validated by signal integration. All variations were detected by PLS-DA, except otherwise stated.

Compounds	**δ [a]/ ppm**	**PLS-DA**		**Variation trend [b]**
		Variation trend	**[VIP] value**	
Acetic	2.07, s	↑	0.49	↑ (< 5%, days 1-7)
Formic	8.23, s	↑	0.61	↑ (< 5%, days 1-18)
GABA	2.51, t	↓	1.73	↓ (6%, days 0-18)
Gallic	7.09, s	↓	1.64	↓ (< 5%, days 2-18)
Higher alcohols [c]	0.86-0.89	↑	0.87	↑ (< 5%, days 1-13)
5-HMF	9.45, s	↑	2.12	↑ (150 %, days 0-18)
Proline [d]	2.10, m	-	-	↓ (< 5%, days 1-18)
Pyruvic	2,40, s	↑	0.70	↑ (< 5%, days 0-7)
Succinic	2.65, s	↑	1.50	↑ (< 5%, days 0-7)
Unassigned 1	1.46, d	↓	2.33	↓ (18%, days 1-13)
Unassigned 3	2.31, d	↑	1.09	↑ (< 5%, days 0-10)
Unassigned 5	3.79, t	↑	2.07	↑ (< 5%, days 1-13)
Unassigned 6	4.09, d	↑	1.93	↑ (13%, days 0-18)
Unassigned 7	6.35, s	↓	1.90	↓ (39%, days 0-18)
Dextrins:				
Av. no. branching points/ molecule [e]	-	↑	-	↑(16%, days 0-18)
Av. no. glucose units/ molecule [e]	-	↑	-	↑ (5%, days 0-7)

[a] Signal chosen within whole spin system, for spectral integration (s, singlet; d, doublet; t, triplet; m, multiplet); [b] Main variation tendency: ↑ or ↓ refer to compounds showing an increasing or decreasing trend with aging and respective variation percentage; [c] Range comprises overlapped signals arising from isobutanol, isopentanol and propanol. [d] Proline identified through 2D correlation analysis alone and not by PLS-DA; [e] Average number of branching points and glucose units per molecule, as determined in previous reports (Duarte *et al.*, 2003).

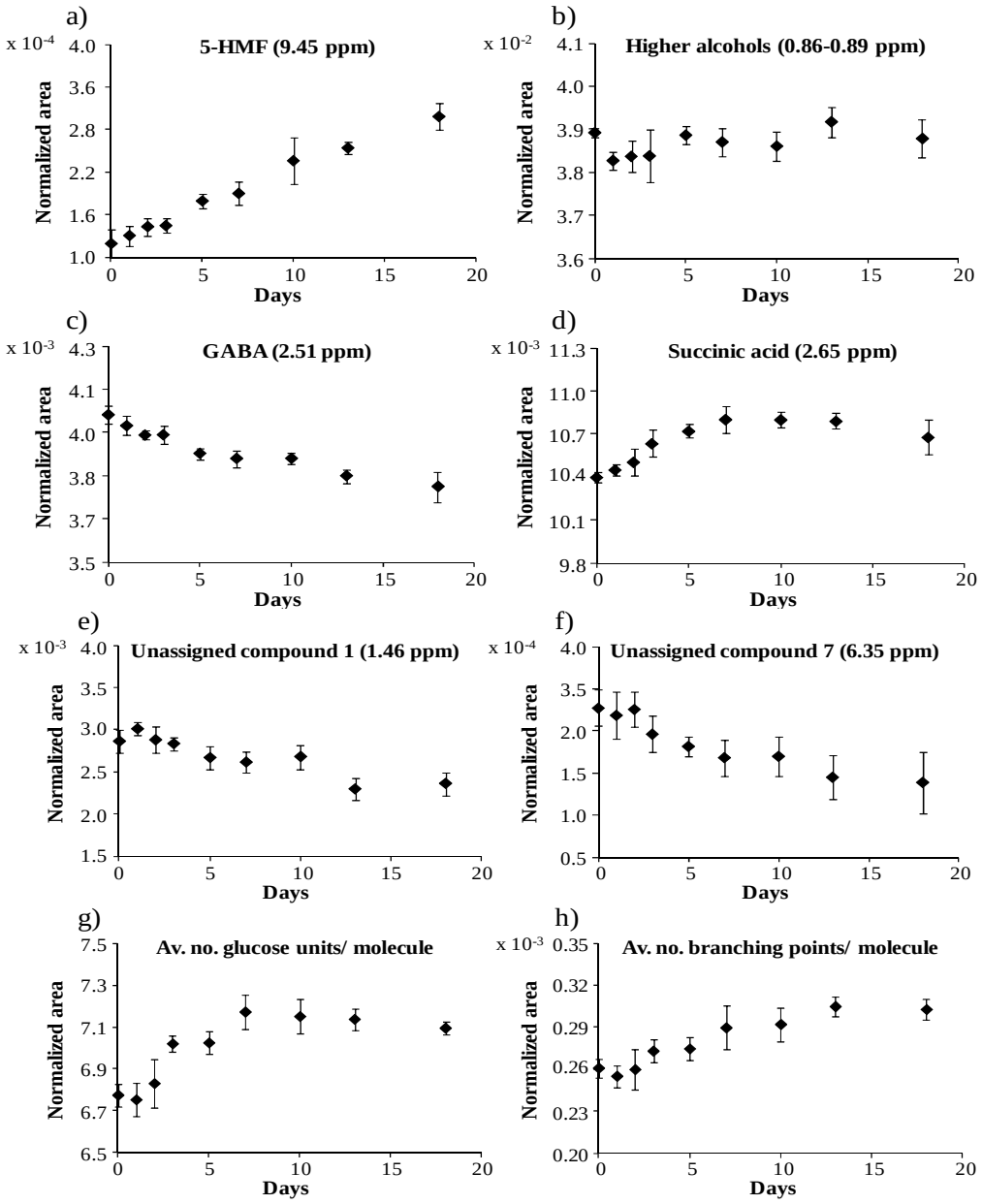

Figure 5.5. Plots of normalized area obtained for some relevant compounds identified as varying as a function of aging days: a) 5-HMF, b) higher alcohols, c) GABA, d) succinic acid, e) unassigned compound 1, f) unassigned compound 7, g) average number of branching points per glucose oligomers molecule and h) average number of glucose units per glucose oligomers molecule.

5-HMF is clearly the most significant aging marker, showing a 150 % variation (through integration at 9.45 ppm) between days 0 and 18 and following an approximately linear increase (Figure 5.5a) as expected. Indeed, as mentioned before, 5-HMF, its formation by Maillard reaction shown in Figure 5.6, is a well known product of sugar degradation, with its increase following an approximately linear increase and

varies logarithmically with the storage temperature (Madigan *et al.*, 1998; Vanderhaegen *et al.*, 2006). Higher alcohols show a very slight (< 5%) overall increasing tendency with aging, with the exception of day 0, which may reflect some hydrolysis of the corresponding esters, as reported for wine (Camara *et al.*, 2006), however, at earlier stages (until day 2) a decrease is noted (Figure 5.5b), possibly reflecting initial oxidative degradation (Vanderhaegen *et al.*, 2006). The small (< 5%) overall increasing tendencies of some organic acids (acetic, formic, pyruvic and succinic) and decrease (- 6%) in GABA (Figure 5.5c) should reflect the ongoing Maillard reactions (Vanderhaegen *et al.*, 2006; Davidek *et al.*, 2006), however, the slight organic acids decrease at days 13 and 18 (Figure 5.5d) suggests that, at this stage, ethanol esterification may determine these compounds content (Vanderhaegen *et al.*, 2006).

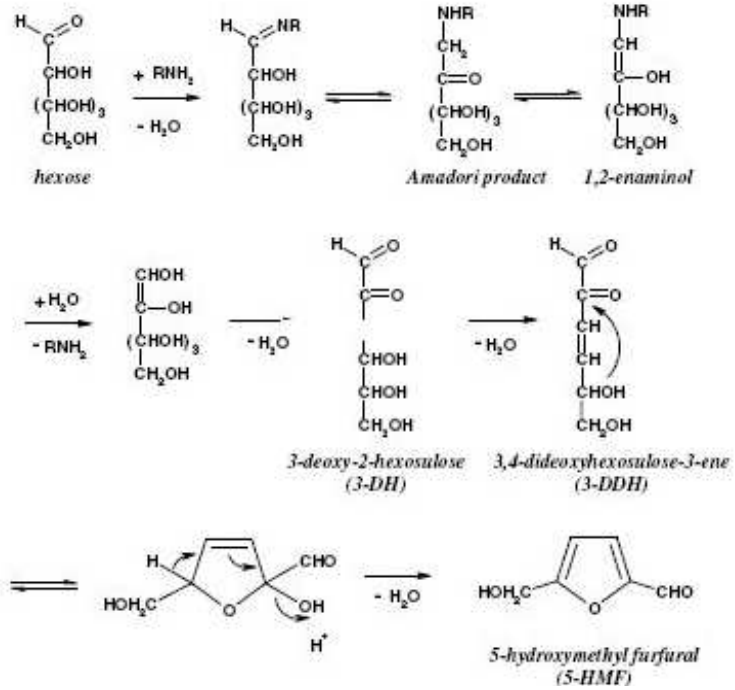

Figure 5.6. Formation of 5-HMF by Maillard reaction (reproduced from Vanderhaegen *et al.*, 2006).

Gallic acid shows a different trend compared to other organic acids, its small decreasing tendency (Table 5.1) possibly reflecting its reaction with reactive oxygen species (ROS) leading to isomerization and formation of polyphenols (Vanderhaegen *et al.*, 2006; Zhao *et al.*, 2010). The observations relating to dextrin changes are themselves interesting, since to our knowledge, no such alterations with beer aging have been reported. Indeed, the relative increases in both the average oligomers chain length

and number of branching points (per molecule) with thermal aging (Figure 5.5g,h and Table 5.1) suggests that dextrin degradation takes place, particularly, after day 3 and possibly to feed Maillard pathways. This seems to occur through a mechanism involving preferentially the shorter and linear oligomers, therefore leading to average increases in chain length and branching degree.

Finally, five still unassigned signals (especially unassigned 1, 6 and 7, with percentage variations of 18, 13 and 39%, respectively) show significant changes with aging (Figure 5.5e,f and Table 5.1), appearing as interesting potential aging markers, particularly after day 3. So far, attempts to identify those unassigned compounds were not conclusive due to their low peak intensity and/or lack of signal coupling that do not allow a deeper understanding of the compound structure, being needed, therefore, to be further studied. An approach to be followed in the future is the use of LC-NMR/MS technique in order to characterize those compounds in question.

5.3.3 Two dimensional (2D) correlation analysis

2D correlation analysis was firstly attempted considering the whole NMR spectra, however, the results obtained reflected only the correlations between the most intense signals, i.e. dextrin resonances, thus not giving useful inter-compound information and masking possible correlations between less intense peaks. Therefore, models were only built for the aliphatic and aromatic regions. Figure 5.7a illustrates a color plot representation of the synchronous 2D correlation map obtained for the aliphatic region, where the highly correlated peaks are identified, either in orange/red (positive) or in blue (negative). For instance, in the 2.3-2.8 ppm region (expansion in Figure 5.7b), the following correlations are recognized for succinic acid (column at 2.65 ppm): 1) positive correlations with acetic acid (under c, overlapped with proline), unassigned 3 (d) and pyruvic acid (e) and 2) negative correlations with higher alcohols (a), unassigned 1 (b), proline (c, overlapped with acetic acid) and GABA (f). This means that succinic acid varies in the same direction as acetic, unassigned 3 and pyruvic (i.e. all increasing with aging), and in the opposite direction of higher alcohols, unassigned 1, proline and GABA (which decrease with aging). The results obtained are listed in Table 5.2 and confirm those obtained by PLS-DA and signal integration, with the exception of higher alcohols, which are suggested by 2D correlation analysis to decrease with aging, contrary to the small increasing tendency seen by PLS-DA and

integration (Table 5.1 and Figure 5.5b). This apparent contradiction probably reflects the low magnitude of the alcohols variation and the effect of the relatively high value found at day 0, so that care should be taken in interpreting the corresponding 2D correlation results.

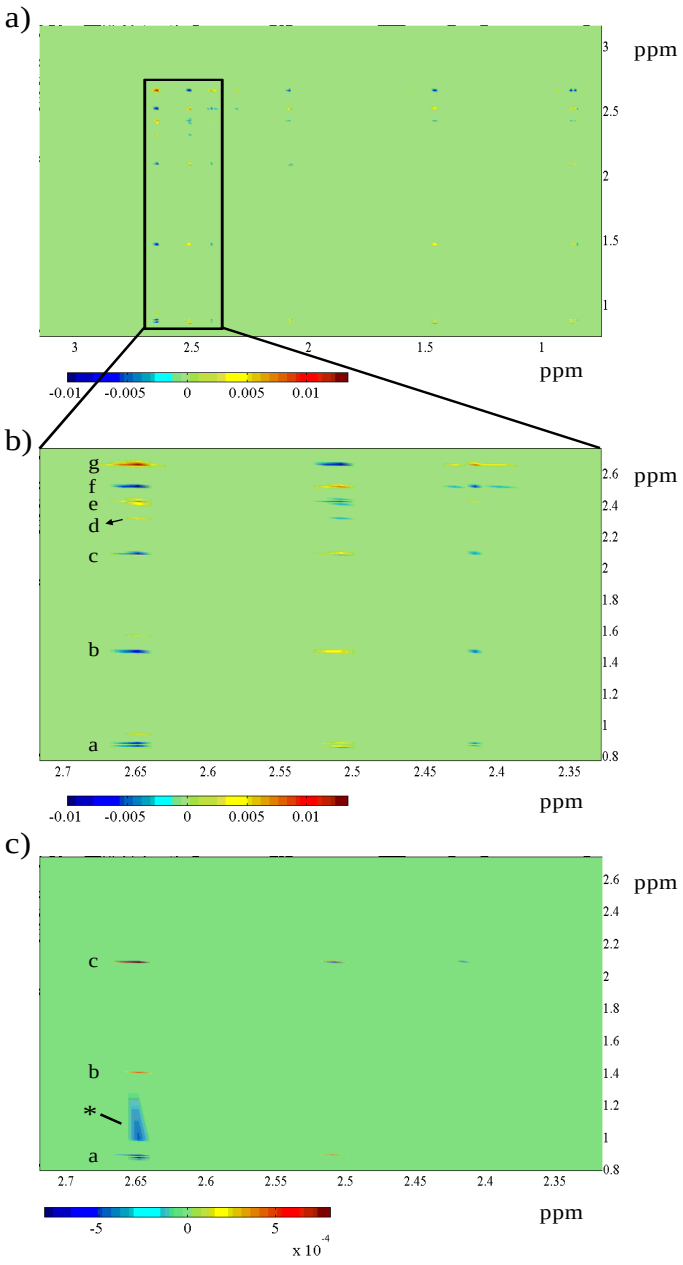

Figure 5.7. Synchronous 2D correlation analysis contour plot obtained for a) aliphatic (0.8-3.1 ppm) region and expansions of b) the 2.4-2.7 ppm; and c) the asynchronous 2D correlation analysis contour plot for 2.4-2.7 ppm. Signals positively and negatively correlated are shown respectively in red and in blue. Compounds correlated with succinic acid were identified as: a, higher alcohols; b, unassigned 1; c, acetic acid and proline; d, unassigned 3; e, pyruvic acid; f, GABA; and g, succinic acid (auto-correlation signal). *: Artefact related to the excluded region of ethanol (1.0-1.3 ppm).

Proline was here identified for the first time in this study as decreasing (< 5%) with beer aging, although not markedly, reflecting the ongoing of Maillard reactions between reducing sugars and amino acids (Vanderhaegen *et al.*, 2006). This observation may have been hindered in PLS-DA due to peak overlap, thus demonstrating the useful complementarity of the 2D correlation analysis in relation to PLS-DA.

As mentioned before, synchronous and asynchronous 2D correlation analysis results provide complementary information. Figure 5.7c shows the expansion of the asynchronous contour map for the 2.3-2.8 ppm region, where succinic acid shows positive correlations with unassigned 1 (b) and acetic acid (under c) and negative correlations with higher alcohols (a) and proline (overlapped with c). Since succinic and acetic acids correlate positively in the synchronous plot (Figure 5.7b), the positive asynchronous correlation expresses a higher (or earlier) variation for succinic, compared to acetic acid. This sign rule is reversed for unassigned 1, higher alcohols and proline, as they are negatively correlated to succinic acid (in the synchronous plot), therefore suggesting a higher (or earlier) variation for unassigned 1 and lesser (or later) variations for alcohols and proline (Table 5.3). Furthermore, the synchronous correlations between a) GABA (↓) and the threesome succinic acid (↑), pyruvic acid (↑) and unassigned 3 (2.31 ppm, broad and weak) (↑) and b) unassigned 7 (singlet at 6.35 ppm) (↓) and 5-HMF (↑) are also noted. The fact that these variations occur to the same extent suggests that they may reflect direct chemical pathways. For instance, the results suggest GABA as a particularly important substrate for Maillard reactions (along with proline and possibly dextrin degradation products), compared to the other amino acids identified by NMR (alanine, leucine, isoleucine, valine, tyrosine and phenylalanine) and for which no relevant consumption was detected. Apparently, GABA consumption is directly correlated to increases in succinic acid, pyruvic acid and unassigned 3. Furthermore, the interesting relationship between 5-HMF and unassigned 7 suggests the latter, still unassigned, compound to be an important 5-HMF precursor. Besides these, other compound correlations are visible (Table 5.2) but their interpretation is hindered by the different variation extensions for each compound, possibly reflecting their involvement in a number of different complex reactions.

Table 5.2. Results of 2D correlation analysis of NMR spectra of beer as a function of aging; (+) and (−): positive and negative correlations. In the interpretation section, compound correlations are described in relation to compound in left column: ↑ or ↓ refer to compounds showing increases or decreases with aging; ***higher***, ***lower*** refer, respectively, to compounds varying to higher, lower extents; ***sync*** refers to compounds varying with the same rate. Ui: unassigned compound i.

	U1	Acetic	Proline	U3	Pyruvic	GABA	Succinic	U7	Gallic	5-HMF
Synchronous results										
U1	-	-	(+)	-	(−)	(+)	(−)	-	-	-
Acetic	-	-	(−)	(+)	(+)	(−)	(+)	-	-	-
Proline	(+)	(−)	-	-	(−)	(+)	(−)	-	-	-
U3	-	-	-	-	-	(−)	(+)	-	-	-
Pyruvic	(−)	-	(−)	-	-	(−)	(+)	-	-	-
GABA	(+)	(−)	(+)	(−)	(−)	-	(−)	-	-	-
Succinic	(−)	(+)	(−)	(+)	(+)	(−)	-	-	-	-
U7	-	-	-	-	-	-	-	-	-	(−)
Gallic	-	-	-	-	-	-	-	-	-	(−)
5-HMF	-	-	-	-	-	-	-	-	(−)	-
Asynchronous results										
U1	-	-	-	-	(−)	(+)	(−)	-	-	-
Acetic	-	-	(−)	-	-	(+)	(−)	-	-	-
Proline	-	(+)	-	-	(+)	(−)	(+)	-	-	-
U3	-	-	-	-	-	-	-	-	-	-
Pyruvic	(+)	-	(−)	-	-	-	-	-	-	-
GABA	(−)	(−)	(+)	-	-	-	-	-	-	-
Succinic	(+)	(+)	(−)	-	-	-	-	-	-	-
U7	-	-	-	-	-	-	-	-	-	-
Gallic	-	-	-	-	-	-	-	-	-	-
5-HMF	-	-	-	-	-	-	-	-	(−)	-

Interpretation (variation direction; variation extent)		
Compound/variation		Related compound variations
U1	↓	GABA: ↓ *lower* ; Pyruvic acid: ↑ *lower* ; Succinic acid: ↑ *lower*
Acetic	↑	GABA: ↓ *higher*; Proline: ↓ *lower* ; Succinic acid: ↑ *higher*
Proline	↓	Acetic acid: ↑ *higher*; GABA: ↓ *higher*, Pyruvic acid: ↑ *higher*; Succinic acid: ↑ *higher*
U3	↑	GABA: ↓ ***sync*** ; Succinic acid: ↑ ***sync***
Pyruvic	↑	GABA: ↓ ***sync***; Proline: ↓ *lower*; Succinic acid: ↑ ***sync***; Unassigned 1: ↓ *higher*
GABA	↓	Acetic acid: ↑ *lower* ; Proline: ↓ *lower*; Pyruvic acid: ↑ ***sync***; Succinic acid: ↑ ***sync***; Unassigned 1: ↓ *higher* ;Unassigned 3: ↑ ***sync***
Succinic	↑	Acetic acid: ↑ *lower* ; GABA: ↓ ***sync*** ; Proline: ↓ *lower*; Pyruvic acid: ↑ ***sync***; Unassigned 1: ↓ *higher* ; Unassigned 3: ↑ ***sync***
U7	↓	5-HMF: ↑ ***sync***
Gallic	↓	5-HMF: ↑ *higher*
5−HMF	↑	Gallic acid: ↓ *lower* ; Unassigned 7: ↓ ***sync***

5.3.4 Conclusions of NMR/MVA analysis

The use of NMR spectroscopy, in tandem with multivariate analysis, enabled the chemical changes occurring during the forced aging (45^{o}C up to 18 days) of beer to be evaluated. PLS-DA and spectral integration analysis showed that a clear aging trend is observed, with fresher beer (up to 3 days of aging) being richer in GABA, gallic acid, linear dextrins and two unassigned compounds (numbered 1 and 7), and poorer in 5-HMF, higher alcohols, acetic, formic, pyruvic and succinic acids, branched dextrins and three other unassigned compounds (named 3, 5 and 6). The following aging stages were characterized by gradually more marked decreases in GABA, unassigned 1 and 7 and ratio linear/branched dextrins and increases in 5-HMF, acetic, pyruvic and succinic acids and unassigned 3 and 6. In addition, over-aged beer (days 13 and 18) shows an inversion of the increasing trend for acetic, pyruvic and succinic acids and unassigned 3. Furthermore, 2D correlation analysis of the NMR data enabled the relevant compound changes to be confirmed in a rapid manner and revealed interesting inter-compound dependences which suggested GABA as a particularly important substrate for Maillard reactions, probably along with dextrin degradation products, with direct impact in the formation of succinic acid, pyruvic acid and unassigned 3. In addition, a still unassigned compound (unassigned 7) was revealed as an apparently important precursor of 5-HMF. The above results show the usefulness of NMR/MVA to detect multiparametric variations of compounds in a complex food mixture such as beer, enabling new insights to be achieved into beer aging chemistry (particularly the relation between beer aging and the ratio linear/branched dextrins; and the inter-dependencies between unidentified peak 3, pyruvic and succinic acids and GABA, and 5-HMF and unidentified peak 7, and unveiling possible new compounds and potential markers of the process.

5.4 Use of GC-MS combined with multivariate analysis (GC-MS/MVA)

In this section, a faster GC-MS analysis methodology was developed and used for identifying/measuring the compositional changes occurring on the complex matrix of a lager beer during it forced aging (at 45^{o}C for up to 18 days). Although extensively employed to identify and quantify aroma/flavour components in beer (Vanderhaegen *et al.*, 2003; Soares da Costa *et al.*, 2004; Saison *et al.*, 2008), the GC-MS technique together with the subsequent process of detection and identification of analytes is time-

consuming, not only due to the chromatographic deconvolution process but also to data annotation and comprehension (Matisova and Domotorova, 2003). Recent advances have demonstrated that the rapid GC analysis procedures, without optimum chromatographic separation (R=1.5), applied to complex foodstuffs (in wine) produces spectral fingerprints containing relevant information, which could be extracted using multivariate analysis (MVA) (Rocha *et al.*, 2006). In this context, faster chromatographic runs were developed, leading to decreased chromatographic resolution, but without losing relevant spectral MS information. Comparison of rapid runs with longer/conventional GC-MS runs was performed in order to validate the findings. GC-MS/PCA enabled the assessment of the main profile changes accompanying beer aging, allowing specific relevant compounds to be identified and deeper knowledge of the aging process to be obtained.

5.4.1 GC-MS characterization of beer: comparison between conventional and rapid runs

To increase the speed of beer analysis, the chromatographic time by conventional GC (41.5 mins) was shown to be reduced by a threefold factor when rapid GC runs are employed (12.8 mins). To confirm that no relevant information loss occurs when rapid GC-MS is performed, chemical characterization of thermally aged beer, with emphasis on the identification of known aging markers, was studied by both GC methodologies. Figure 5.8 shows typical chromatograms obtained for a lager beer using a) conventional and b) rapid GC runs. Overall, the same 25 compounds were identified for both methodologies (Table 5.3), the two chromatograms showing similar profiles, although a slight loss of chromatographic separation was observed on the chromatogram corresponding to rapid runs (Figure 5.8b), with the convolution of peaks 7 and 8 (corresponding to furfuryl alcohol and valeric acid), 17 and 18 (2,3-dihydro-3,5-dihydroxy-6-methyl-4(H)-pyran-4-one (DDMP) and capric acid) and 23 and 24 (tetrahydro-4-hydroxy-4-methyl-2(H)-pyran-2-one and benzeneacetic acid). Examples of characteristic mass spectra obtained for a) 5-HMF and b) caproic acids; and c) structures of identify compounds are shown in Figure 5.9.

To ensure that the rapid GC methodology also maintained the (semi-) quantitative ability of the conventional GC runs to characterize the detected compound, comparison of the normalized areas obtained by rapid and conventional GC-MS data for the relevant

compounds in beer aging process was performed, using the signal of the internal standard (3-octanol) as intensity reference. The compound peaks choice was based on their intensity, contribution to the aging process and the occurrence of convolution in their peaks on rapid GC runs. Accordingly, the following compounds were chosen and integrated: acetic acid, furfural, 2,3-butanediol, butyric acid, phenylacetaldehyde, furfuryl alcohol, valeric acid, 2-phenylethyl acetate, caproic acid, caprylic acid, vinylguaiacol, DDMP, capric acid, diethyl succinate, 5-HMF, maltoxazine, tetrahydro-4-hydroxy-4-methyl-2(H)-pyran-2-one, benzeneacetic acid and a unknown compound, named unknown 1, with a corresponding *m/z* values of 43.

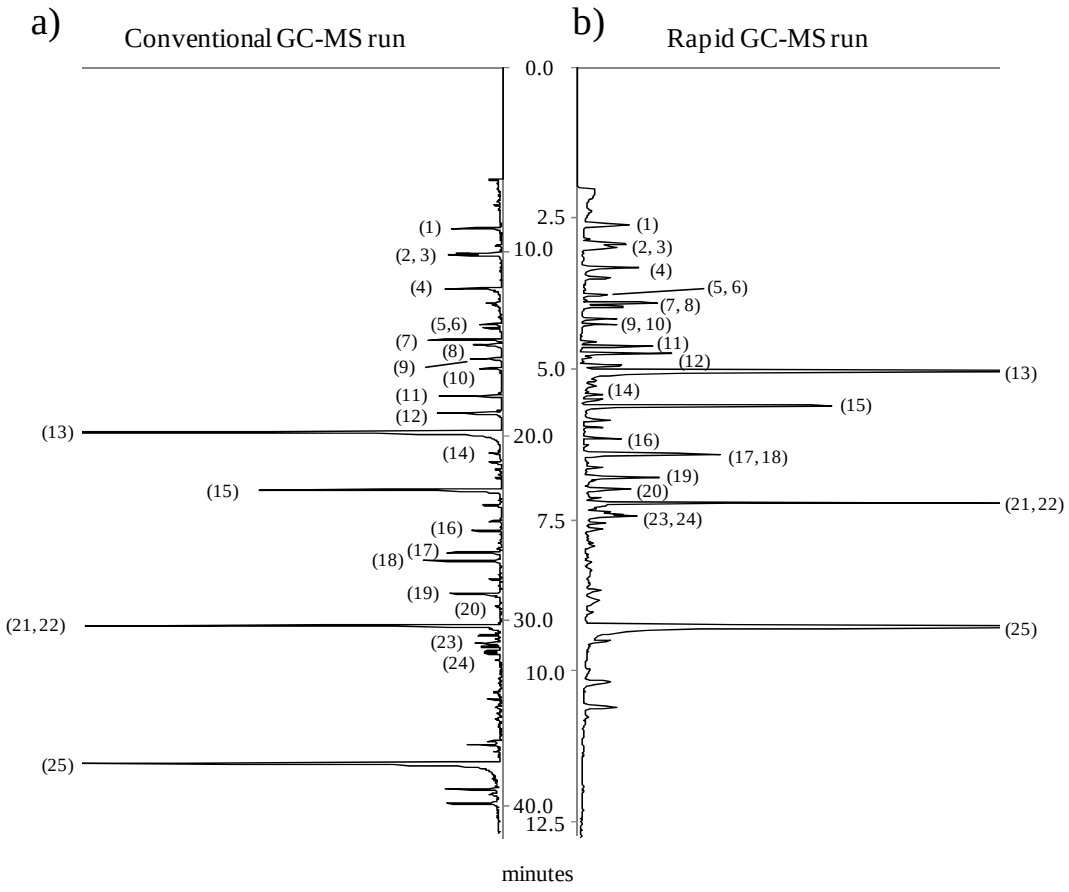

Figure 5.8. Chromatograms obtained for aged beer at day 18 by a) conventional GC-MS and b) rapid GC-MS runs. The main signals were assigned as: (1) 3-octanol, internal standard, (2) acetic acid, (3) furfural, (4) 2,3-butanediol, (5) butyric acid, (6) phenylacetaldehyde, (7) furfuryl alcohol, (8) valeric acid, (9) methionol, (10) unknown peak 1 (characteristic *m/z* 43), (11) 2-phenylethyl acetate, (12) Caproic acid, (13) β-phenylethanol, (14) 2-acetylpyrrole, (15) caprylic acid, (16) vinylguaiacol, (17) DDMP, (18) capric acid, (19) diethyl succinate, (20) benzoic acid, (21) 5-HMF, (22) maltoxazine, (23) tetrahydro-4-hydroxy-4-methyl-2(H)-pyran-2-one, (24) benzeneacetic acid and (25) tyrosol.

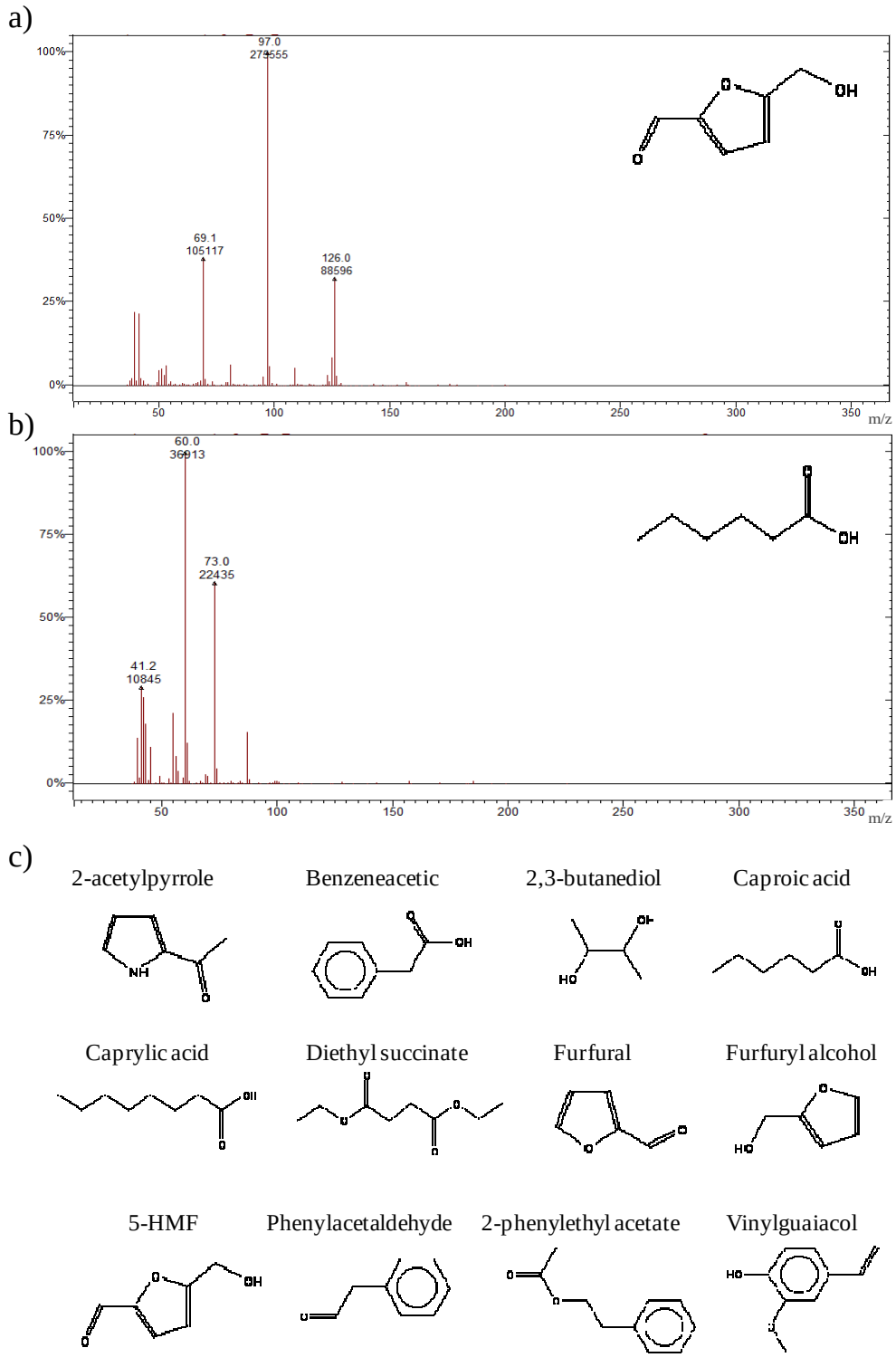

Figure 5.9. Mass spectra and corresponding structures of) 5-HMF and b) caproic acid; c) structures of some compounds found in aged beer.

Table 5.3. Structural information corresponding to the spots identified on the GC-MS run. Retention times indicated for the identified compounds refer to long (left) and rapid (right) GC runs. The underlined values represent the *m/z* values used for peak integration.

Retention time (mins)	Main *m/z* fragments	Compounds
8.8/ 2.6	41/ 55/ 83	3-octanol (internal standard)
10.1/ 2.9	43/ 60	Acetic acid
10.2 / 3.0	39/ 67/ 95	Furfural
12.0/ 3.3	45/ 57	2,3-butanediol
14.1/ 3.9	42/ 60/ 73	Butyric acid
14.1/ 3.9	91/ 120	Phenylacetaldehyde
14.8/ 3.9	41/ 69/ 98	Furfuryl alcohol
15.2/ 4.0	41/ 60/ 73	Valeric acid
15.9/ 4.2	61/ 73/ 106	Methionol
16.4/ 4.3	43/ 61/ 75	Unknown 1
17.9/ 4.6	43/ 104	2-phenylethyl acetate
18.8/ 4.8	60/ 73/ 87	Caproic acid
20.0/ 5.1	91/ 122	β-phenylethanol
21.0/ 5.3	94/ 109	2-acetylpyrrole
23.0/ 5.6	60/ 73/ 101/ 115	Caprylic acid
25.2/ 6.1	107/ 135/ 150	Vinylguaiacol
26.3/ 6.4	43/ 101/ 144	2,3-dihydro-3,5-dihydroxy-6-methyl-4(H)-pyran-4-one (DDMP)
26.9/ 6.5	60/ 73/ 129	Capric acid
28.6/ 6.8	73/ 101/ 128	Diethyl succinate
29.4/ 7.0	77/ 105/ 122	Benzoic acid
30.5/ 7.2	69/ 97/ 126	5-hydroxymethyl furfural (5-HMF)
30.6/ 7.3	165/ 179	Maltoxazine
31.0/ 7.4	43/ 71	Tetrahydro-4-hydroxy-4-methyl-2(H)-pyran-2-one
31.2/ 7.5	91/136	Benzeneacetic acid
38.0/ 9.3	77/ 107/ 138	Tyrosol

Comparing the integral response of the selected peaks as a function of the days of aging, for both conventional and rapid GC runs, it became apparent that the response was similar for the majority of the studied compounds, with variations lower than 20%, although the quantitative requirements of a minimum acceptable chromatographic resolution of R = 1.5 (Stafford, 2003) were not fulfilled. Examples of this comparison

are shown in Figure 5.10 for: a) caprylic acid, b) diethyl succinate, c) DDMP, d) furfural, e) 5-HMF, and f) vinylguaiacol. One exception was noted involving 5-HMF, which showed variations of approximately 30% at days 13 and 18 of thermal aging (figure 5.9e), with conventional runs giving higher values, a fact for which no explanation could be advanced at this stage. In any case, apart from 5-HMF at later aging stages, it became apparent that no relevant information is lost when the rapid GC-MS methodology approach is employed.

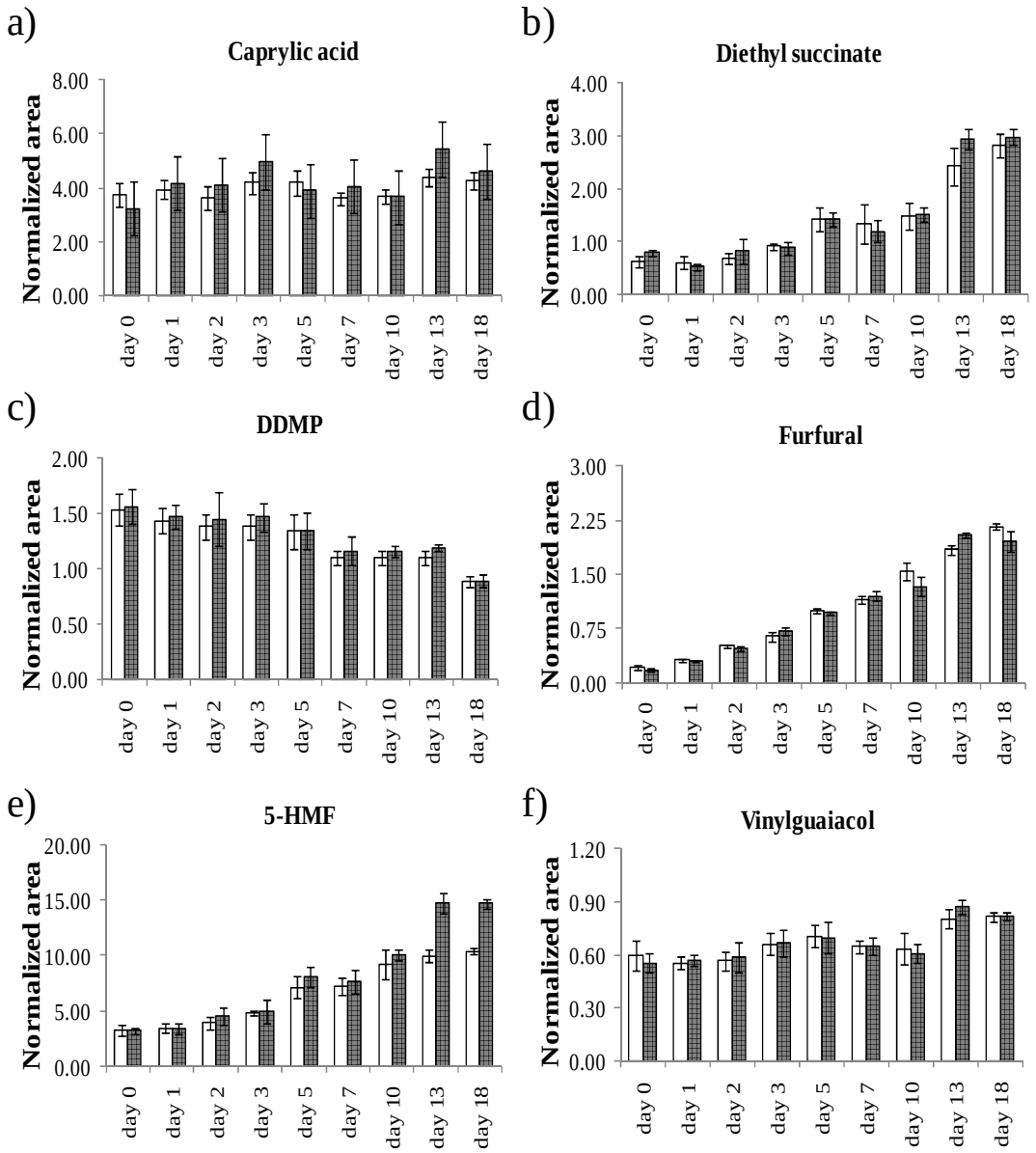

Figure 5.10. Example of normalized areas plots corresponding to the characteristic peaks of a) furfural (extraction of *m/z* value 95) and b) 5-HMF (extraction of *m/z* value 97), as a function of deterioration time, obtained for conventional (dark bars) and rapid GC-MS runs (blank bars).

5.4.2 Multivariate analysis of GC-MS data (GC-MS/MVA) as a function of aging

PCA models were built based on rapid and conventional GC-MS data to aid the identification of the compositional variations occurring during beer aging. Figure 5.11a shows the results obtained for rapid GC-MS data, indicating that a clear trend exists, as a function of aging extent, with PC1 accounting for 44.6% of the total variability. PC1 was the only component that captured the variability related to the aging extension, with the remaining components showing no meaningful patterns. A similar result was obtained for conventional GC runs (Figure 5.11b), accounting for 37.8% of the total variability. Plotting both PC1 scores results (obtained for conventional and rapid GC methods) against each other, a correlation with r = 0.92 was obtained (not shown), thus demonstrating the similarity between the two datasets. Therefore, the following results will only describe the rapid GC-MS/MVA approach.

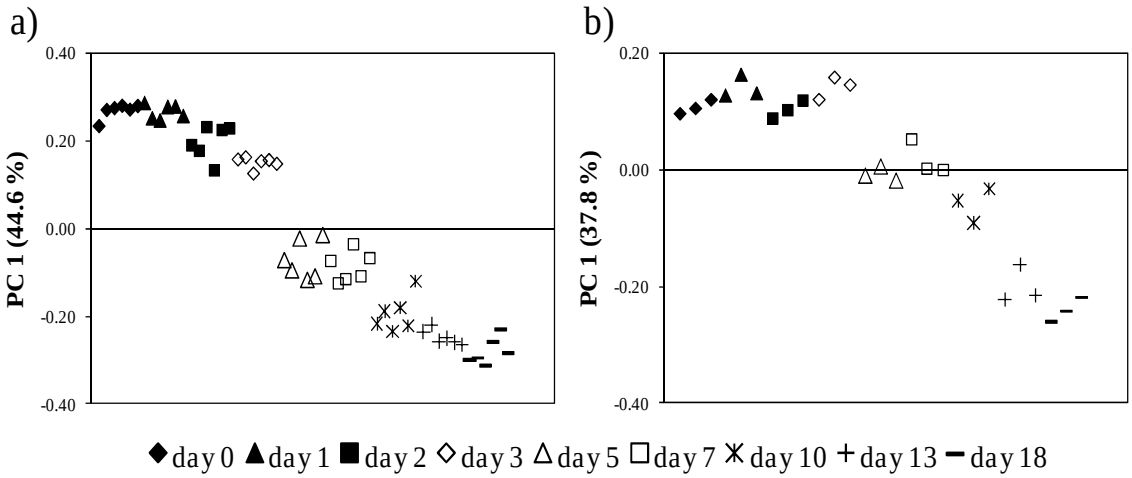

Figure 5.11. PC1 scores plot of a) rapid and b) conventional GC-MS data obtained for the monitoring of the thermal deterioration process.

In Figure 5.11a, fresher beer samples are located in positive PC1, whereas more aged samples are positioned in negative PC1. To interpret this trend, the corresponding contour PC1 loadings plot was obtained (Figure 5.12). This plot combines the chromatographic data corresponding to each beer sample with the *m/z* fragmentation data of each scan, allowing the identification of the main *m/z* fragments responsible for the trend observed in Figure 5.11a. Negative spots (in blue and yellow in figure 5.12) are related to *m/z* fragments which increase in concentration during aging. On the other hand, positive spots (in red) correspond to *m/z* fragments decreasing with the aging

process. Overall, 12 spots were identified by rapid GC-MS/MVA as varying with beer aging, namely acetic acid, 2,3-butanediol, caproic acid, caprylic acid, diethyl succinate, DDMP, furfural, 5-HMF, maltoxazine, phenylacetaldehyde, 2-phenylethyl acetate and unknown compound 1 (*m/z* 43). The interpretation of such spots is shown in Table 5.4, where one still unidentified relevant compound (named unknown 1) is noted. PCA results were partially validated by integration of the identified peaks and examples of the compounds variations, as a function of days of thermal deterioration, are shown in Figure 5.13 for a) 5-HMF, b) diethyl succinate, c) furfural, d) phenylacetaldehyde, e) DDMP, f) maltoxazine, and g) unknown compound 1.

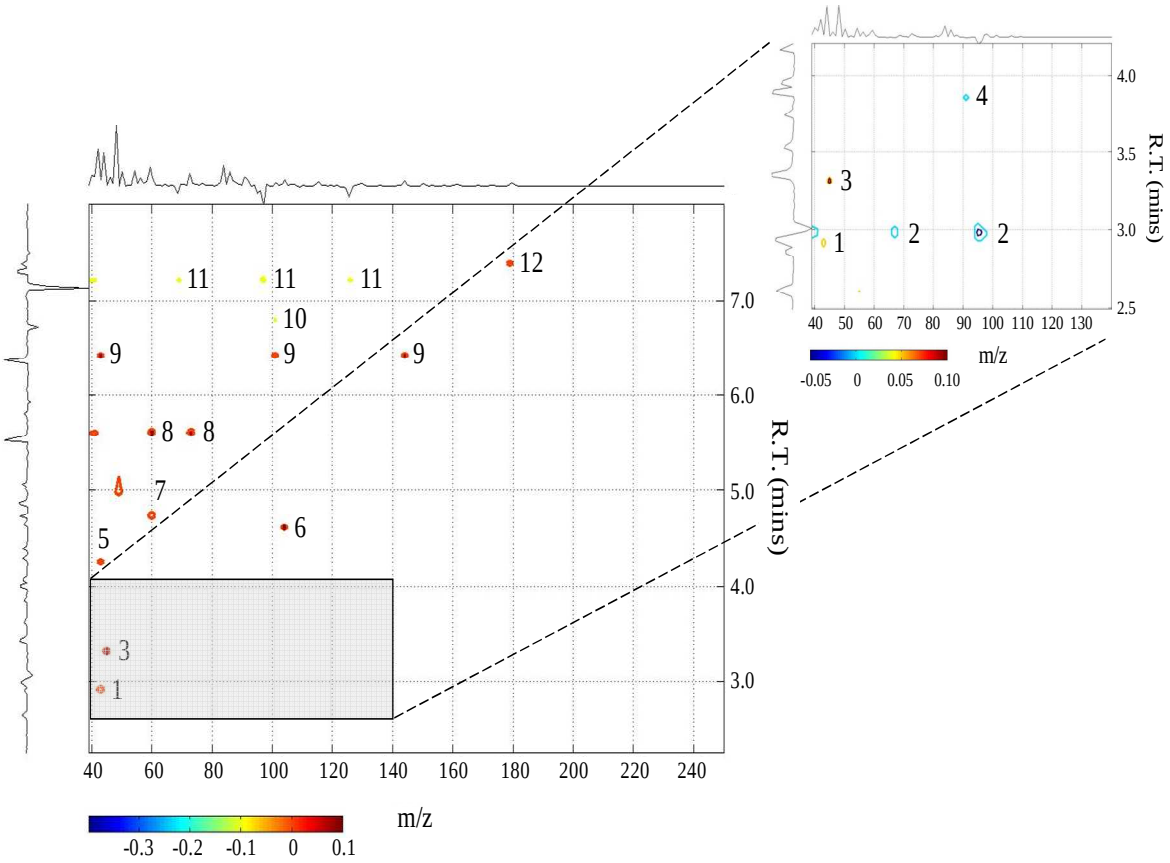

Figure 5.12. PC1 contour loadings plot obtained for rapid GC-MS data, combining chromatographic data and *m/z* fragmentation data corresponding to each scan. The inset shows the expansion of the region between 2.5 – 4.2 mins and *m/z* fragments range between 39 and 140. Compounds identified as varying are 1) acetic acid, 2) furfural, 3) 2,3-butanediol, 4) phenylacetaldehyde 5) unknown compound 1 (*m/z* 43), 6) 2-phenylethyl acetate, 7) caproic acid, 8) caprylic acid, 9), DDMP, 10) diethyl succinate, 11) 5-HMF, 12) maltoxazine. In blue and yellow (negative values): spots increasing with aging process; in red (positive values): spots illustrated as decreasing with aging.

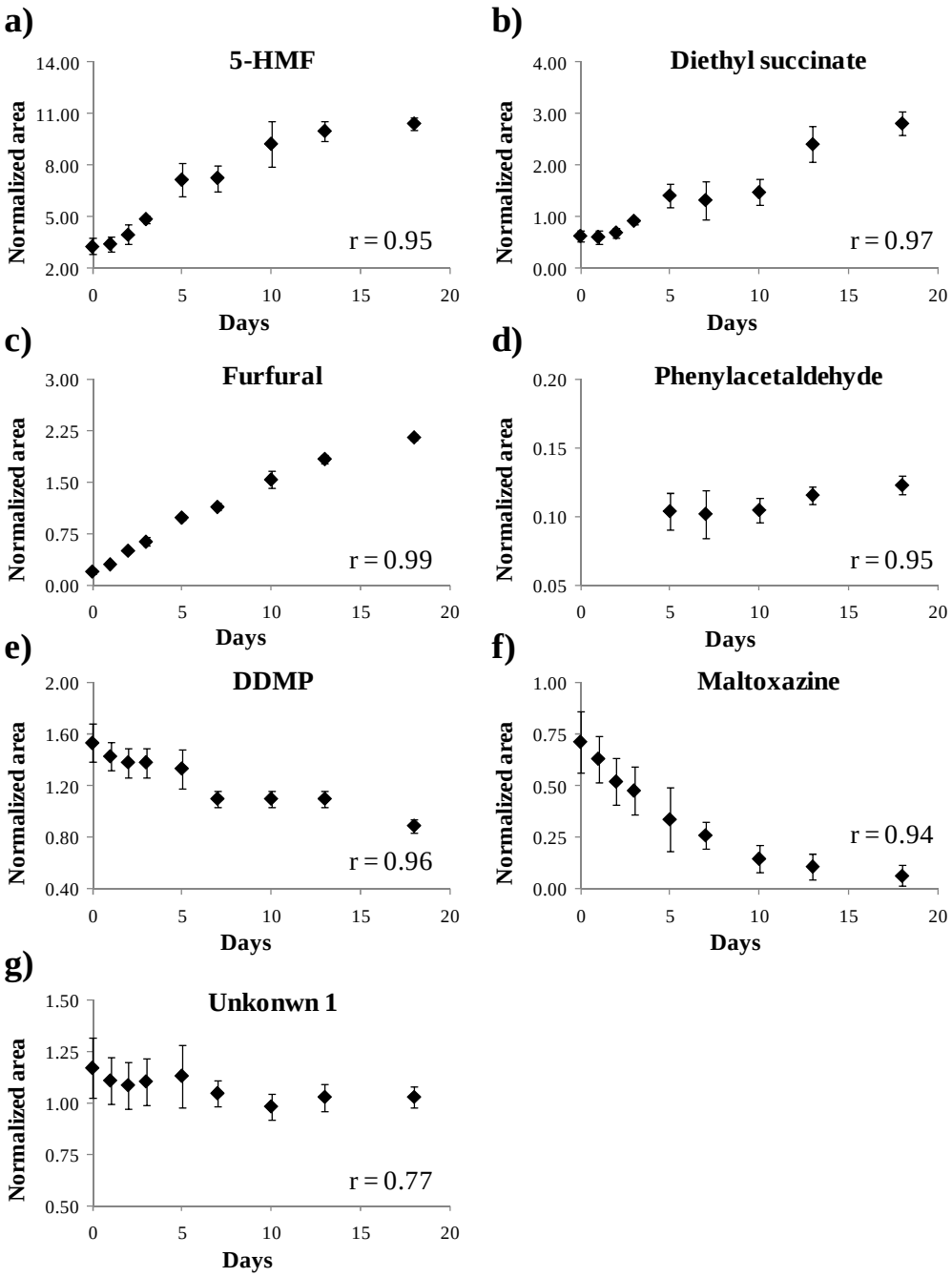

Figure 5.13. Plots of normalized areas as a function of days of aging obtained for some of the compounds identified on PC1 as responsible for deterioration, namely a) 5-HMF, b) diethyl succinate, c) furfural, d) phenylacetaldehyde, e) DDMP, f) maltoxazine, and g) unknown 1. Correlation coefficients, r, are shown in each plot.

The spots corresponding to 5-HMF (*m/z* 69, 97 and 126), diethyl succinate (*m/z* 101), furfural (*m/z* 67 and 95) and phenylacetaldehyde (*m/z* 91) showed negative intensity, confirming their expected increasing tendency with thermal deterioration process. Accordingly, approximate linear increases were detected for 5-HMF, diethyl

succinate and furfural, with correlation values r > 0.95 and variation percentages between fresh (day 0) and deteriorated beer in day 18 of 218%, 946% and 353%, respectively. Indeed, 5-HMF (already identified in NMR/MVA) and furfural are well known products of sugar degradation (Madigan *et al.*, 1998; Vanderhaegen *et al.*, 2006). The obtained variation of 5-HMF as a function of aging confirms the trend observed by NMR/MVA, which also identifies 5-HMF as a major compound increasing with forced aging with a % variation between fresh (day 0) and aged beer (day 18) of 157% and 218 % for NMR and rapid GC-MS, respectively. The difference between the degrees of variation between NMR and GC-MS techniques may be justified by difficulties in the integration of NMR signal in the earlier stages of aging (until day 2), since 5-HMF signal was very weak.

Diethyl succinate, a product of the esterification of ethanol with succinic acid, has also been shown to increase linearly, its formation rate being dependent on storage temperature (Vanderhaegen *et al.*, 2003). An increasing trend was also noted for phenylacetaldehyde, although this compound was found present in trace amounts. This compound is a Strecker aldehyde, derived from phenylalanine degradation, and is a known aging marker, its formation being greatly affected by temperature and level of dissolved oxygen (Soares da Costa *et al.*, 2004). Due to phenylacetaldehyde low content, especially at the earlier days of aging, peak integration was only possible after day 5 of forced aging (figure 5.13d).

The positive intensity spots identified comprised, amongst others, DDMP (*m/z* 43, 101 and 144), maltoxazine (*m/z* 179) and unknown compound 1 (*m/z* 43), meaning that they are more abundant in fresh beer and, therefore, decreasing with aging. In fact, those compounds show a clear decrease, with variation percentages between fresh beer (day 0) and aged beer at day 18 of 42%, 91% and 11% for DDMP, maltoxazine and unknown 1, respectively. For DDMP and maltoxazine, an approximately linear decrease with days of aging was observed, with r values > 0.94. DDMP, originating from hexoses degradation, has been observed as an important intermediate of Maillard reactions, being sensitive to higher temperatures (Kim and Baltes, 1996). Maltoxazine, a compound formed by Maillard reactions involving proline (Yaylayan and Mandeville, 1994) and present in barley malt and beer (Tressl *et al.*, 1982; Yaylayan and Mandeville, 1994), to our knowledge, as not been previously related to the aging

process. Although a trend was observed for unknown compound 1, with r 0.77, variations were relatively small compared to the previous compounds.

Table 5.4. Structural information corresponding to the spots identified on the PC1 contour loadings plot. The underlined values represent the *m/z* values used for peak integration and the variation percentage (% variation) refers to compounds variation between day 0 (fresh beer) and day 18 of aging (last day of forced aging).

Spot number	Retention time (mins)	Main *m/z* fragments	Compounds	Variation trend	% Variation
1	2.9	43	Acetic acid*	-	-
2	3.0	67/ 95	Furfural	Increase	946
3	3.3	45	2,3-butanediol*	-	-
4	3.9	91	Phenylacetaldehyde	Increase	28
5	4.3	43	Unknown 1	Decrease	11
6	4.6	104	2-phenylethyl acetate*	-	-
7	4.8	60	Caproic acid*	-	-
8	5.6	60	Caprylic acid*	-	-
9	6.4	43/ 101/ 144	DDMP	Decrease	42
10	6.8	101	Diethyl succinate	Increase	353
11	7.2	69/ 97/ 126	5-HMF	Increase	218
12	7.3	150/ 179	Maltoxazine	Decrease	91

* Compounds for which no clear trend was identified by analysis of their normalized areas as a function of thermal degradation.

In addition, 2-phenylethyl acetate (*m/z* 104), 2,3-butanediol (*m/z* 45), acetic (*m/z* 43), caproic (*m/z* 60) and caprylic acids (*m/z* 60) were also identified as positive intensity spots on the PC1 contour loadings plot (decreasing with aging). However, by analysis of their integrals as a function of the aging period, no statistically relevant trends were identified for these compounds, thus indicating that these compounds may not be adequate markers of beer aging (at least under the specific experimental conditions used in this work). The importance and involvement of these compounds in beer aging has, however, been reported (Diaz-Maroto *et al.*, 2005; Vanderhaegen *et al.*, 2006; Vanderhaegen *et al.*, 2007). For instance, 2-phenylethyl acetate seems to tend to hydrolyze with aging, especially at higher temperatures (Vanderhaegen *et al.*, 2007). In addition, 2,3-butanediol can react with aldehydes (acetalization of aldehydes) producing

cyclic acetals (Vanderhaegen *et al.*, 2006) and acetic acid can be involved in various reactions during storage, such as the chemical hydrolysis of acetate esters (Vanderhaegen *et al.*, 2007) and the oxidation of ethanol (Vanderhaegen *et al.*, 2006). Finally, caproic and caprylic acids are medium linear fatty acids that, during beer storage, can react with alcohols producing the corresponding esters (Diaz-Maroto *et al.*, 2005).

5.4.3 Conclusions of GC-MS/MVA analysis

In this thesis a reliable and fast GC-MS/MVA methodology to monitor beer aging process was developed, not only reducing chromatographic separation time, without losing relevant spectral information and the (semi-) quantitative ability to characterize the identified compounds, but also improving data interpretation by using a MVA approach to analyze GC-MS data.

The time needed for chromatographic deconvolution was reduced by threefold compared to conventional GC runs and, although some loss of resolution was detected, equivalent compositional information was observed for both rapid and the conventional GC-MS runs. The subsequent GC-MS/MVA methodology enabled the identification of the degree of aging of beer and the identification of the specific compounds related to the process. In fact, the forced aging process of beer was clearly correlated to the increase of the well known aging compounds 5-HMF, furfural, diethyl succinate and phenylacetaldehyde. Furthermore, a decrease in DDMP and maltoxazine were observed, together with one still unknown compound (unknown 1). To our knowledge, this variation in DDMP and maltoxazine was here observed, for the first time, as correlated with beer deterioration, while the remaining unassigned compound may, in time, unveil a possible new aging marker.

5.5 Use of MIR spectroscopy combined with multivariate analysis (MIR/MVA)

In this section, the use of MIR spectral data in combination with MVA was performed for the rapid assessment of the degree of beer aging. MIR spectroscopy shows great potential for rapid analysis of beer as it provides precise measurements in a fast and nondestructive manner, with simple sample preparation (only degassing is necessary). Moreover, MIR is an ‘ease to use’ technique, considerably less expensive

compared to NMR and GC-MS methods. On the other hand, IR methods have inherent difficulties on the analysis of complex mixtures and aqueous solutions (major influence of water bands), giving rise to spectra that, when compared with NMR spectroscopy, show lesser qualitative information about beer chemical composition and the changes occurring during the aging process.

In a first approach, direct analysis by MIR spectroscopy was applied to study the aging process of beer. Due to the major influence of the characteristic absorptions bands of water in the direct beer spectra, a new methodology has been applied: the droplet evaporation of a beer samples in MIR. This methodology consists in placing a droplet of beer in the MIR crystal, accompanying the evaporation process and acquiring the corresponding spectrum. PCA and PLS-DA methods were applied to the MIR data corresponding to each studied methodology in order to extract, in a rapid manner, the useful information about the aging process.

5.5.1 Multivariate analysis of direct MIR spectra

Figure 5.14a shows the total MIR spectrum and the expansion of the fingerprint region (1300-850 cm^{-1}) for fresh (day 0), in black, and aged beer at day 18, in blue. By visual observation, no differences between the two spectra are identified (completely overlapped). Furthermore, water dominates the beer spectrum, with absorption bands located at 3200 and 1600 cm^{-1}, corresponding to OH stretching and OH bending vibrations, respectively. In the fingerprint region, the most intense bands appear at 1151, 1084, 1045, 1030 and 998 cm^{-1}, reflecting the major contributions of beer carbohydrates and ethanol, as seen by comparison with reference spectra shown in Figure 5.15b. Indeed, ethanol characteristic bands are located at 1087 and 1045 cm^{-1} reflecting C-O stretching bands. Carbohydrates content is dominated mainly by branched and linear dextrins. To mimic linear and branched dextrins spectra, suspensions of amylose (linear molecule made up of α(1→4) bound glucose molecules) and amylopectin (branched molecules constituted by glucose units linked in a linear way by α(1→4) and branched by α(1→6) glycosidic bonds), respectively, were analyzed. Amylose spectrum shows absorption bands at 1151, 1080 and 1010 cm^{-1}, whereas amylopectin shows bands at 1151, 1085 and 1021cm^{-1}, with the corresponding bands arising from C-O and C-C stretching vibrations, C-OH bending vibrations and CH related modes. Since no visual differences were detected between samples from day

0 and 18 (Figure 5.14a), the use of MVA procedures to identify possible compositional variations with beer aging process may be useful.

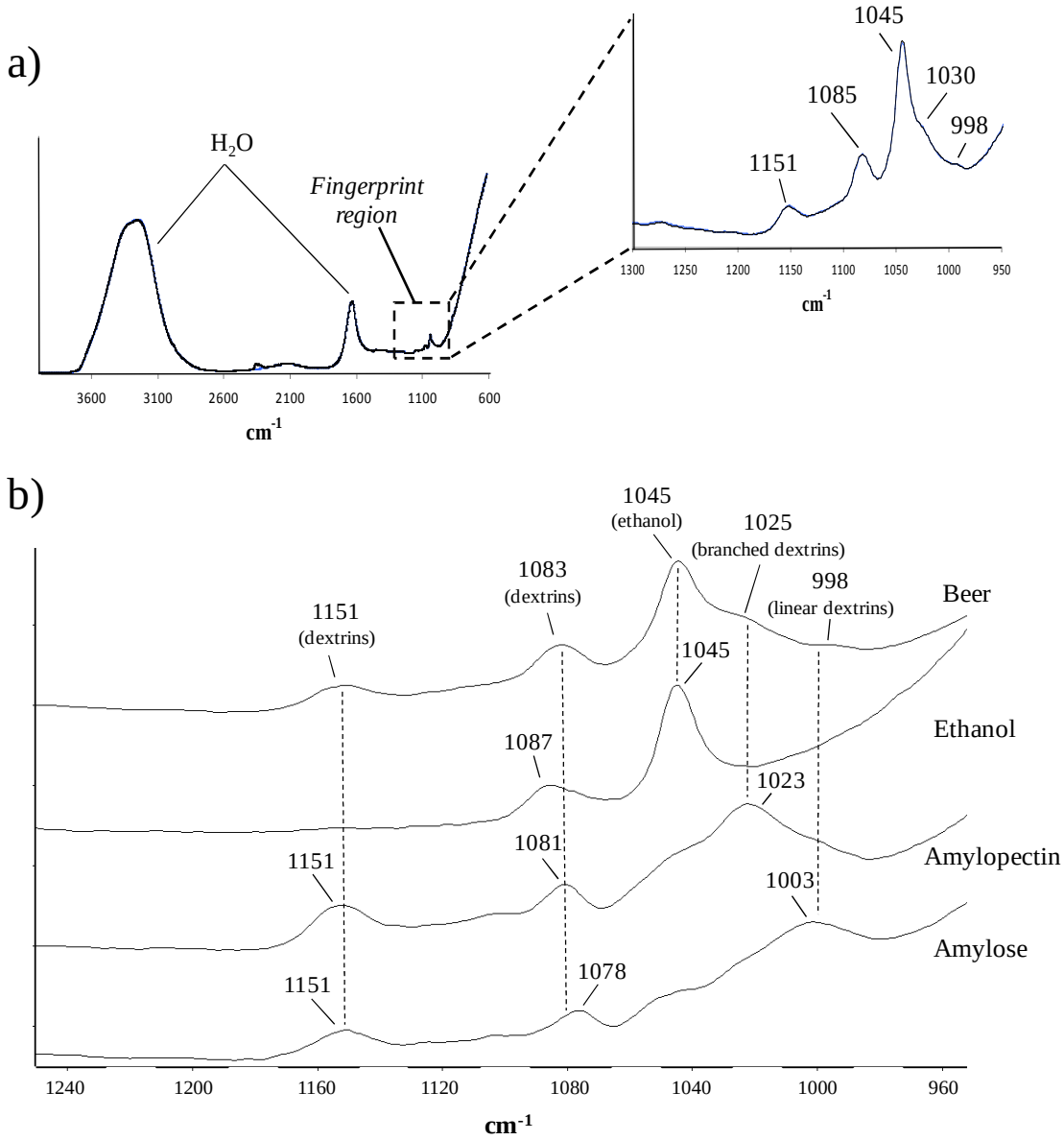

Figure 5.14. Direct MIR spectra of a) fresh beer (day 0), in black, and aged beer (day 18), in blue, with expansion of the fingerprint region (1300-900 cm-1); and b) comparison of beer, ethanol solution, amylopectin and amylose suspensions (1250-950 cm^{-1} spectral region).

PCA and PLS-DA models of the MIR spectra between 1300-850 cm^{-1} and corresponding to the days of aging were performed, using two data processing methods: autoscaling and second derivative (followed by autoscaling). Both PCA and PLS-DA models of the direct MIR data using autoscaling processing showed a separation

between the days of aging as seen on the corresponding 3D representations (Figure 5.15a), with the identification of four sub-groups: 1) days 0, 1 and 5; 2) days 2 and 3; 3) days 7 and 10; and 4) days 13 and 18. Comparing the PCA and PLS-DA scores plots, a higher dispersion between samples from the same day of aging is detected using PCA method. This may reflect the fact that PCA is an unsupervised method and, therefore, no prior information about the relationship between MIR spectra and days of aging is given to the built up of the MVA method. Although the PLS-DA scores plot shows lower sample dispersion and a better separation between days of aging, it is characterized by a negative Q^2 value of -0.03, indicating that the model has no predictive power

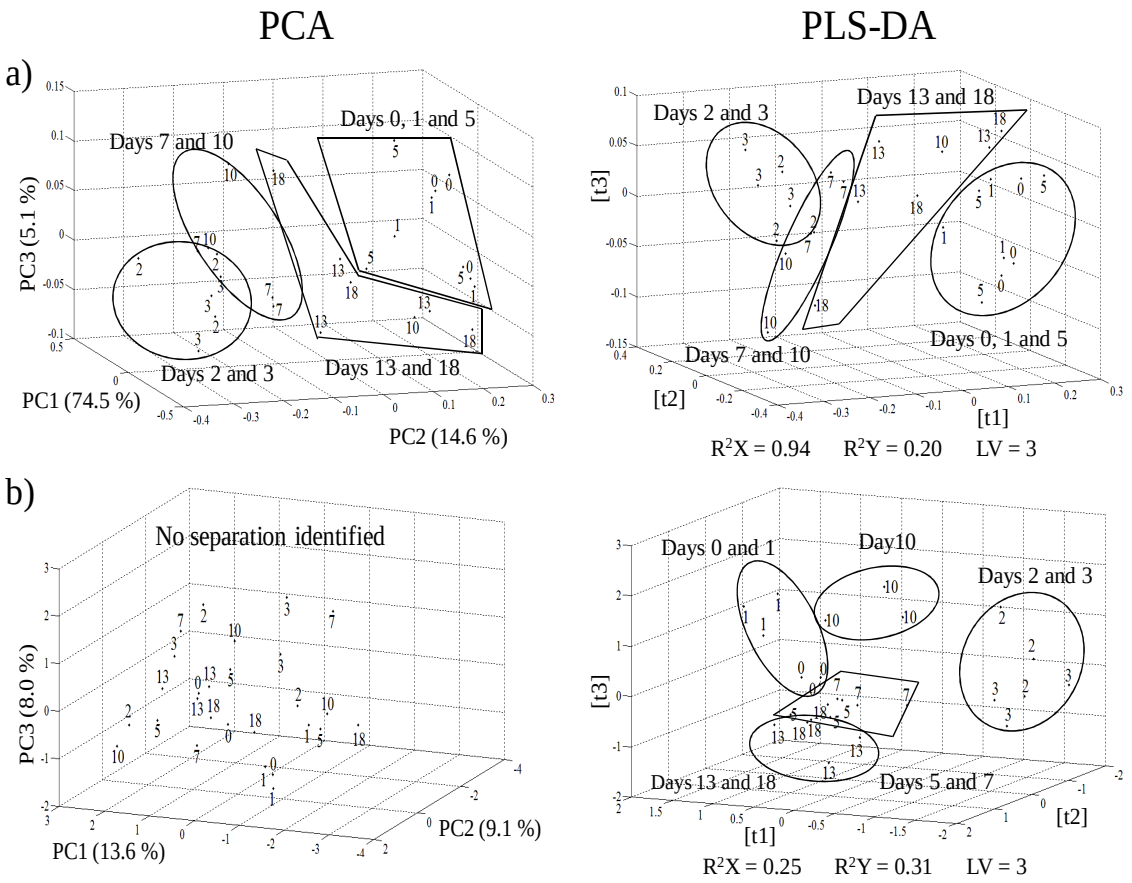

Figure 5.15. PCA, in the left, and PLS-DA, in the right, 3D scores scatter plots of direct MIR spectra (region between 1300-850 cm^{-1}) with a) autoscale and b) second derivative/ autoscale treatment of the data. Numbers indicate aging days and grouping shapes were drawn to guide the eye.

Regarding the PCA and PLS-DA models using a second derivative processing (Figure 5.15b), the results showed that in the PCA method no separation trend is detected, whereas in the PLS-DA model a distinction between the different days of

aging is identified, showing five sub-groups: 1) days 0 and 1, 2) days 2 and 3, 3) days 5 and 7, 4) day 10; and 5) days 13 and 18. These results confirm the observation expressed before, that the supervised method of PLS-DA allows a clearer clustering between the samples corresponding to each day of aging, although the model is once again non-validated, Q^2 (-0.05). Overall, the PLS-DA method using second derivative processing show a clearer grouping between days of aging, although non-validated. Indeed, when using the autoscaling processing, the same tendency of separation was obtained for PLS-DA models and PCA methods, whereas using the second derivative processing beers corresponding to days 5 and 7 group together and day 10 beers (grouped with day 7 in autoscaling models) form a new sub-group. Individual analysis of each latent variable or principal component, in the PLS-DA and PCA, respectively, was performed. However, no significant trend of separation in each plot was identified. Therefore, the corresponding loadings weights plots were not able to give relevant information about the similarities and differences between aged beer samples.

Although the MVA methods applied to direct MIR data were able to distinguish between the days of aging, no clear aging degree trend was identified.

5.5.2 Multivariate analysis of droplet evaporation MIR spectra

Figure 5.16a shows the evolution of the MIR spectrum by applying the droplet evaporation methodology. When the droplet starts to evaporate, a decrease of the water bands (located at 3200 and 1600 cm^{-1}) as well as the enhancement of the signals located in the regions between 3000-2800 cm^{-1} and 1500-800 cm^{-1} is noted. From 20 minutes onwards, the spectral features remained constant, indicating that the evaporation process has stabilized using the following spectra in the built up of the MVA models.

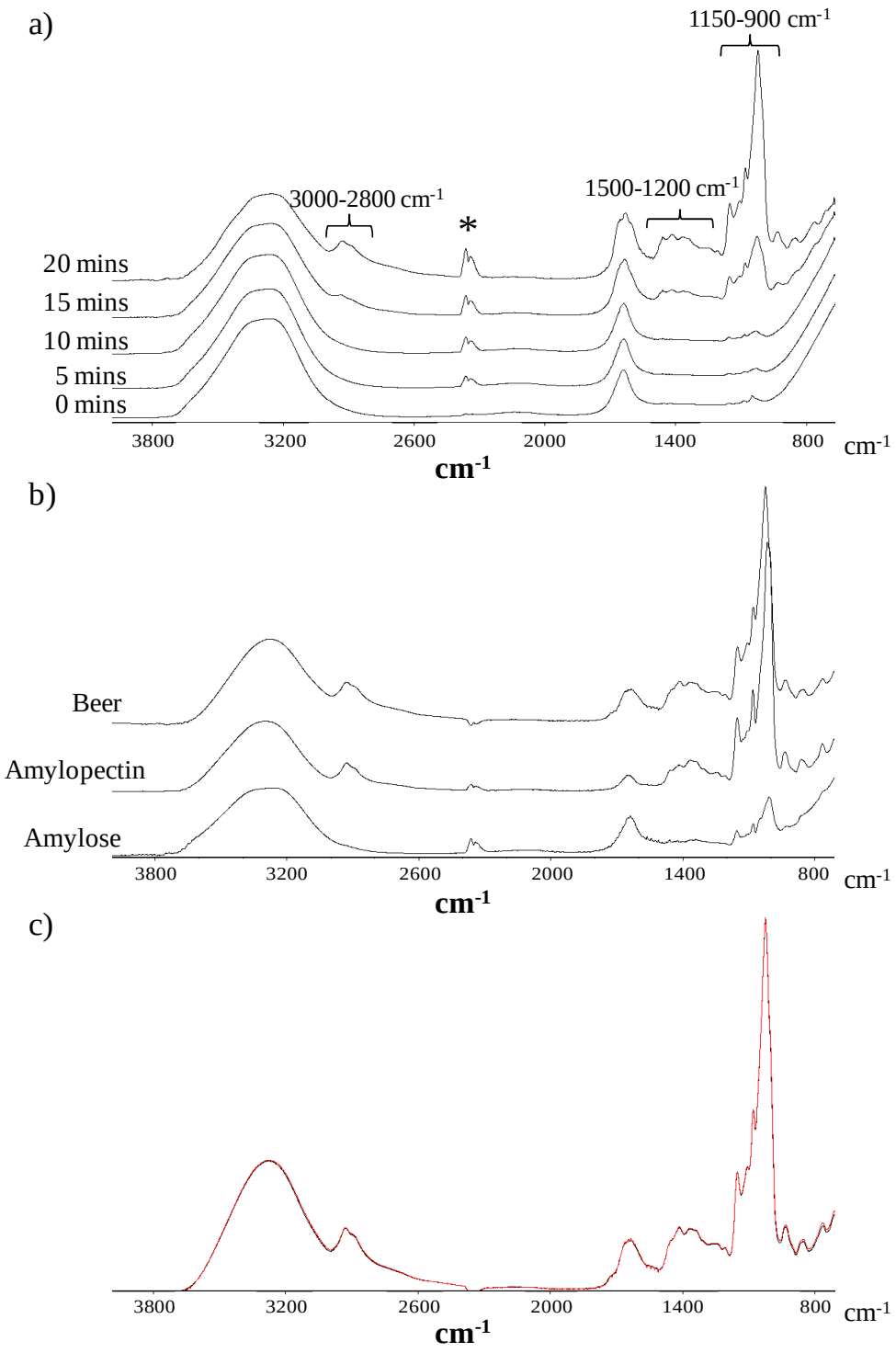

Figure 5.16. MIR spectra using the droplet evaporation methodology: a) series of spectra recorded during the beer droplet evaporation; b) spectra of beer, amylopectin and amylose suspensions; and c) spectra of fresh beer (day 0), in black, and aged beer (day 18), in red. * band arising from CO_2(C=O asymmetric stretching).

Comparing the characteristic beer droplet evaporation spectrum with those corresponding to amylose and amylopectin (Figure 5.16b), it is seen that the characteristic bands of carbohydrates dominate the beer spectrum, even when applying the evaporation methodology. Indeed, beer droplet evaporation spectrum is similar to those obtained for amylopectin (branched polymer), being characterized by bands arising between 3000-2800 cm^{-1} (related with CH stretching vibrations), between 1500-1200 cm^{-1} (arising from vibrations related to CH and COOH groups) and at 1150, 1105, 1075, 1020 and 1000 cm^{-1} (already mentioned in the direct MIR analysis, and corresponding to C-O and C-C stretching vibrations, C-OH bending vibrations and CH related modes). By visual comparison between the spectra corresponding to fresh beer (day 0, in black) and to aged beer (day 18, in red) shown in Figure 5.17c, no relevant intensity variations between the spectra were identified, as seen when using the direct MIR methodology.

PCA and PLS-DA models of the droplet evaporation MIR data were built, with basis on two spectral regions: i) 3000-850 cm^{-1} and ii) 1500-850 cm^{-1}. Both autoscaling and second derivative data processing were applied, with the models corresponding to autoscaling processing showing a separation between days of aging, whereas those related with the second derivative processing showed no separation between days (higher dispersion between samples from the same days) and, therefore, only the MVA models using autoscaling processing are shown and discussed.

Using the MIR spectra region between 3000-800 cm^{-1}, the 3D PCA scores plot (PC1 vs. PC2 vs. PC3, which together accounts for 98.6% of the total variability) showed a distinction between aging days, with the identification of several sub-groups, namely: days 0 and 1, day 2, day 3, days 5 and 7, day 10, day 13 and day 18 (Figure 5.17a, left). The corresponding PLS-DA model, characterized by R^2X and R^2Y values of 0.99 and 0.24, respectively, show a similar separation between days, with the exception of day 10 which is grouped with day 18 (Figure 5.18a, right). Analyzing either each PC or latent variables plots individually, no separation tendency between days of aging was detected, therefore, the corresponding loadings weights plots could not be used to obtain relevant information about which bands reflect the similarities and differences between samples (as seen when using the direct MIR methodology).

Regarding the 3D PCA scores scatter plot obtained for the spectral region between 1500-850 cm^{-1} (Figure 5.18b, left), the three main PC's accounting to 99.5% of the total

variability, a similar grouping between days was identified when compared to the PCA scores plots obtained for 3800-850 cm^{-1} region. The corresponding representation of the PLS-DA scores plot, characterized by R^2X and R^2Y values of, respectively, 0.99 and 0.24, showed a higher dispersion between samples from the same day of aging, with the identification of four sub-groups: 1) days 0, 1 and 5; 2) days 3 and 7; 3) day 13; and 4) days 2 and 18. However, similarly to the observed on the MVA models corresponding to 3800-850 cm^{-1} MIR data, no clear separation tendency between the days of aging was identified by individual analysis of the main PC or latent variable scores plot.

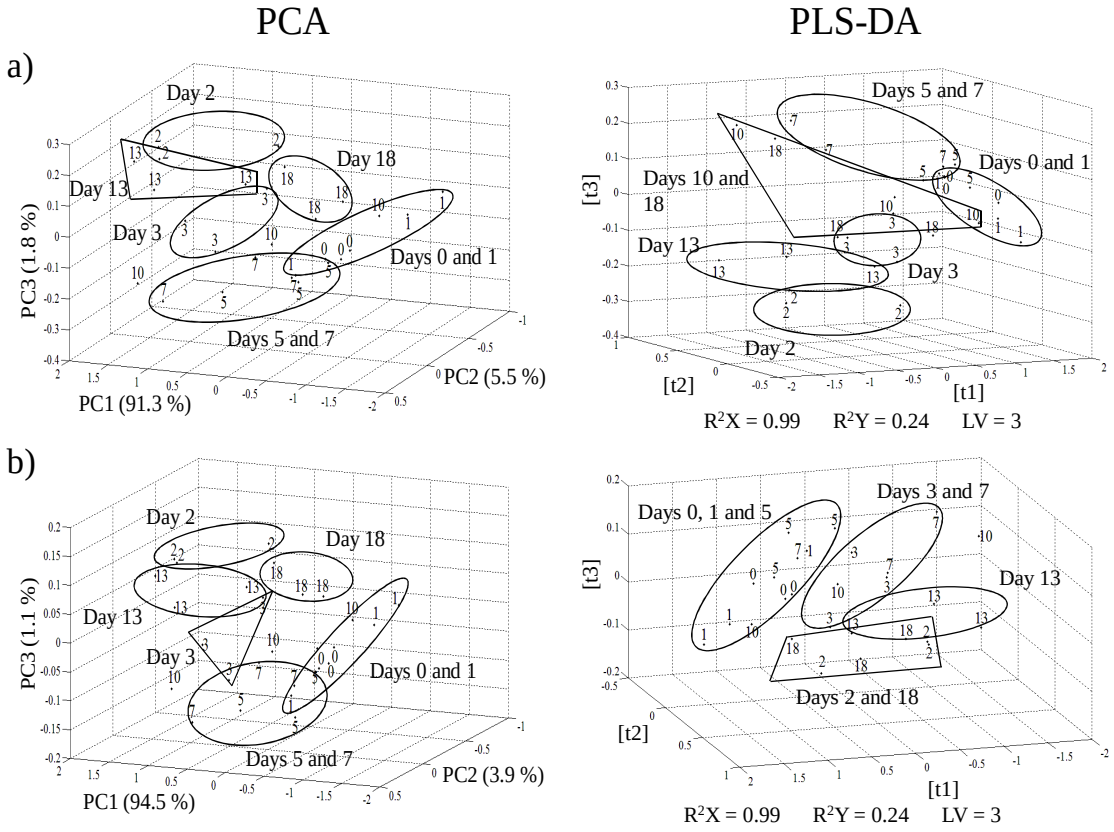

Figure 5.17. 3D PCA (left) and PLS-DA (right) scores scatter plots of droplet evaporation MIR data using autoscaling processing for a) 3000-800 cm^{-1} and b) 1500-800 cm^{-1} regions. Numbers indicate aging days and grouping shapes were drawn to guide the eye.

5.5.3 Conclusions of MIR/MVA analysis

The use of MIR in tandem with multivariate analysis, MIR/MVA, was attempted to accompany the forced aging process of beer. Firstly, the direct MIR analysis approach was followed, with the corresponding spectra showing to be dominated by the water signals, although signals corresponding to branched and linear dextrins and

ethanol can also be detected. PCA and PLS-DA models were applied to the direct MIR data, enabling the identification of a separation within days of aging, although no clear aging degree trend was identified. Since the direct MIR spectrum is clearly dominated by the water signals, an alternative methodology consisted on the evaporation of a beer droplet and subsequent spectral acquisition was applied. This methodology allowed the decrease of the water bands and subsequent enhancement of the remaining signals, particularly branched dextrins. MIR/MVA methods were applied for the spectral regions between 3000-800 cm^{-1} and 1500-800 cm^{-1}. As in the case of direct MIR, the MVA models were only able to distinguish beers between studied days, however no consistent aging trend was detected.

Although the beer droplet evaporation methodology seemed to be a good approach to overcome the water influence limitations encountered on the direct MIR spectra, low reproducibility of the evaporation process was detected and, therefore, developments and/or optimization are required if application of this methodology is to be envisaged. Examples of possible future improvements include a more rigorous control of room temperature and humidity (although controlled, still some significant fluctuations occur: room temperature, 23 (± 3) ºC, and humidity 35 (± 8) %), which, when studying very similar samples as in this case, can originate a higher spectral variability than those originated by the real compositional differences.

5.6 Data correlation for NMR/GC-MS, sensory/NMR and sensory/GC-MS data sets

In this section, the application of MVA methods allowing data correlation between the different analytical domains employed on the accompanying of the beer aging process (i.e. NMR and GC-MS data) and their correlation with beer sensory properties are described. Since MIR data showed some limitations on the identification of an aging degree tendency and needed further optimization, MIR data was not studied in this section.

Principal component transform-outer product analysis-partial least squares-discriminant analysis (PCT-OPA-PLS-DA) and statistical heterospectroscopy (SHY) were applied to NMR and GC-MS data sets, enabling the identification of relationships and/or correlations between compositional changes occurring in different quantitative domains as a function of the days of aging. OPA is a method of data fusion which

enables the combination of different domains, in this case NMR and GC-MS data, by mutual weighting of all the signals/variables, making it possible to detect co-evolutions in signals acquired in those two domains using MVA methods, such as PCA and PLS-DA (Barros *et al.*, 1997; Jaillais *et al.*, 2005; Barros *et al.*, 2008). On the other hand, SHY between NMR and GC-MS data consists on the calculation of the linear correlations between signal intensities for each variable of NMR data with the ones corresponding to GC-MS domain, enabling the identification of signals of the same molecule as well as metabolites whose concentrations are inter-dependent.

Finally, an exploratory study about the potential predictive ability of NMR and GC-MS data, employing the quantitative method PLS1 regression models, to assess the organoleptic characteristics of beer during the forced aging process was also performed, taking into account the possible relationships between either the NMR or GC-MS variables with the response obtained by sensory analysis.

5.6.1 Principal component transform-outer product analysis-partial least squares-discriminant analysis (PCT-OPA-PLS-DA) in NMR and GC-MS data fusion

Prior to data fusion between NMR and GC-MS using OPA methodology, principal component transform (PCT) was applied to each data set, NMR (866 points x 54 spectra) and GC-MS (57200 points x 54 runs), in order to decompose each data set into principal components (PC's) corresponding to the most important sources of variability (54 x 54).

PLS-DA was then applied to the resulting outer product (OP) matrix to detect the relations between the two domains. The results obtained indicated a clear aging degree tendency in the main latent variable, [**t1**], with values of R^2X and R^2Y of 0.83 and 0.35, respectively, and a Q^2 of 0.26 (Figure 5.18). This aging tendency occur in a stepwise manner, with fresher beer (until day 3) being located in positive [**t1**], whereas the more aged ones (day 5 onwards) in negative [**t1**]. The resulting scores plot shows a very similar separation between days of aging when compared to the one obtained in the GC-MS/MVA (section 5.4), thus suggesting a stronger influence of the GC-MS data on the obtained OP matrix. This behavior can be explained by the fact that the intensities variations detected in GC-MS domain are significantly higher than those observed in

NMR, therefore, the influence of GC-MS signals in the resulting OP matrix is also higher. This is observed on the variation percentages values between day 0 (fresh beer) and beer at day 18 (last day of aging) of the compounds identified as varying, for each analytical technique. While NMR data show slight compositional variations between fresh and aged beer (compounds variation lower than 40%, except for 5-HMF), in GC-MS data compositional variations higher than 100% were observed for several compounds (e.g. 5-HMF, diethyl succinate and furfural).

The interpretation of the corresponding loadings weights (**w1**) plots is complicated since the obtained OP matrix corresponds to a cube. Thus, for each NMR variable (bucket, ppm) a corresponding 2D contour plot expressing the related GC-MS information (retention time vs. *m/z* values) is obtained. Some examples of the resulting loadings 2D contour plots are shown in Figure 5.19 for a) 5-HMF, b) succinic acid, c) GABA, d) H1α red., e) unassigned 1, and f) unassigned 7. The identified spots correspond to the variables in the OP matrix (NMR and GC-MS domains) that show higher intensities, with their color reflecting either increasing (blue and yellow, negative values) or decreasing (brown and red, positive values) variations with the aging process.

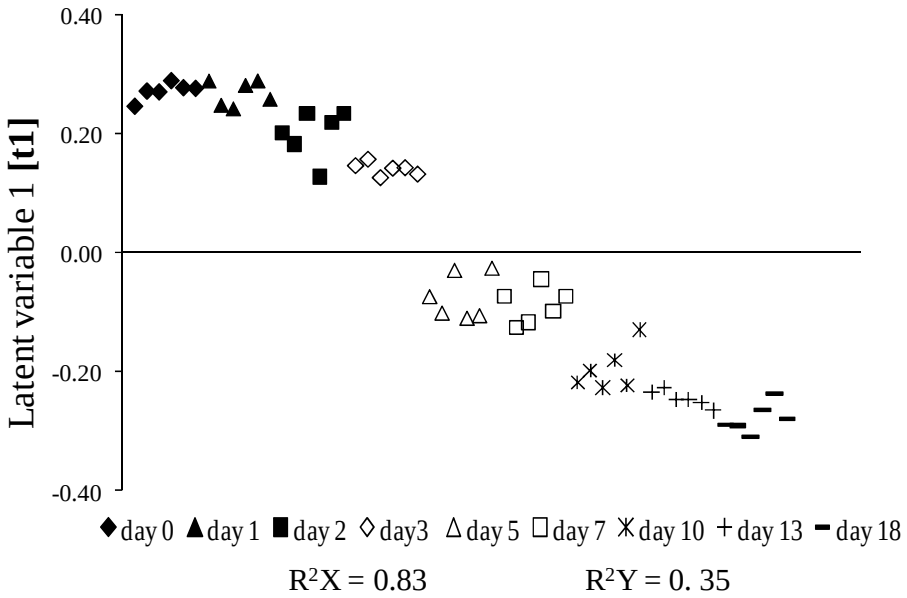

Figure 5.18. Latent variable 1, [**t1**], scores plot obtained from outer-product (OP) matrix between NMR and GC-MS data, as a function of beer forced aging.

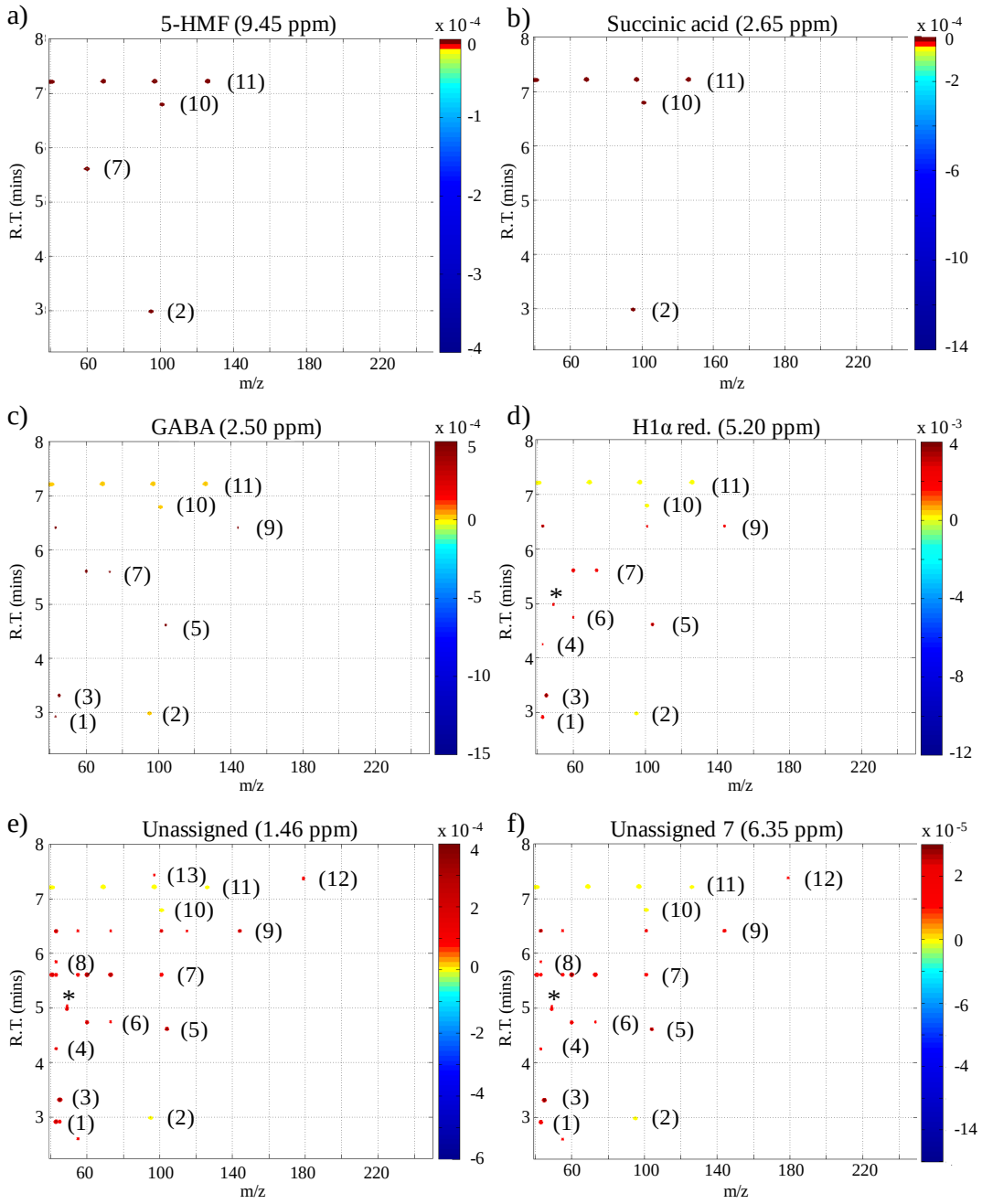

Figure 5.19. Loadings weights (**w1**) contour plots obtained by PCT-OPA-PLS-DA, between NMR and GC-MS (retention time vs. *m/z* fragments) data and corresponding to particular NMR signals a) 5-HMF, b) succinic acid, c) GABA, d) H1α red., e) unassigned 1, and f) unknown compound 7. Compounds identified as varying in GC-MS domain are 1) acetic acid, 2) furfural, 3) 2,3-butanediol, 4) unknown compound 1 (*m/z* 43), 5) 2-phenylethyl acetate, 6) caproic acid, 7) caprylic acid, 8) possible ketone (*m/z* 43 and 114), 9), DDMP, 10) diethyl succinate, 11) 5-HMF, 12) maltoxazine, and 13) benzeneacetic acid. Negative values (yellow and blue): spots increasing with the aging; and positive values (brown and red): spots decreasing with aging. * Artefact related to the solvent (dichloromethane).

Considering all 866 loadings (**w1**) 2D contour plots (corresponding to the 866 variables of the NMR data), it is seen that, although some GC-MS spots vary between the 2D contour plots, the spots corresponding to 2,3-butanediol, 2-phenylethyl acetate and DDMP (characterized as decreasing with aging by GC-MS) and 5-HMF, diethyl succinate and furfural (present in higher quantities in the last days of aging by GC-MS) are identified in all plots as relevant compounds, thus indicating the highest weight of GC-MS in the OP matrix, mentioned before. The only exceptions are seen for the NMR variables: 5-HMF, succinic acid (Figures 5.19a,b, respectively) and branched dextrin H1α (1→6) resonance (increase with aging identified by NMR), in which only the GC-MS spots 5-HMF, diethyl succinate and furfural (all increasing with aging by GC-MS) are identified as relevant. This reflects the fact that when the intensities in the two domains are simultaneously high, the corresponding OP variable will be, consequently, higher than the remaining variables. Furthermore, benzeneacetic acid ($C_8H_8O_2$) and a possible ketone (tentative assignment, 3.85 mins, *m/z* values 43 and 114) (spots from GC-MS domain) were identified as decreasing with aging process. However, by analysis of their integrals as a function of the aging period, no statistically relevant trend was identified.

Other data pre-treatment were tested (namely unit variance scaling, UV) for the development of PLS-DA (not shown), however no aging tendency was identified in that model. Overall, the PCT-OPA-PLS-DA method was able to identify the same compounds varying with aging as those detected by individual NMR-MVA and/or GC-MS/MVA. Although no relevant new additional information was obtained, these results suggest that PCT-OPA-PLS-DA can be an important tool for metabonomic studies, as it allows the data fusion of two analytical techniques using a compressed space (the Principal Component Transform framework), avoiding the use of very wide matrices, without significant loss of information and showing potential to identify relationships between variables of both domains (NMR and GC-MS) in order to characterize more deply a specific variation (in this case the aging process of beer).

5.6.2 Statistical heterospectroscopy (SHY) between NMR and GC-MS data

Statistical heterospectroscopy (SHY) was applied to correlate the NMR and GC-MS data sets in order to identify consistent linear correlations between the major

compounds (viewed by NMR) and the lower content compounds (viewed by GC-MS), many of which known degradation markers.

The correlation coefficients matrix built between NMR and GC-MS data shows the dimensions of a cube, meaning that, as seen previously for the OPA-PLS-DA method, for each NMR variable (chemical shift) a corresponding 2D correlation contour plot with the dimensions of GC-MS data (retention time vs. *m/z* values) is obtained. The correlation scale varies from -1 to +1, where -1 indicates perfect anticorrelation (or negatively correlated) between NMR and GC-MS data, +1 indicates perfect correlation, and zero corresponds to pairs of variables with no correlation. To enhance the extractable information from the correlation plots, highlightening only the more relevant correlations, a correlation coefficient cutoff of ± 0.80 was impose. Furthermore to reduce the chances of observing false correlations, all correlations with a p value > 0.001 were set to zero (Maher *et al.*, 2011).

Using the correlation coefficients cutoff at ± 0.80, the NMR resonances corresponding to 5-HMF, anomeric protons (H1) for β red., α red., α (1→6) and α (1→4) in branched and linear oligomers (all related with the ratio linear/branched dextrins), GABA, succinic acid and unassigned compounds 1, 6 and a new signal, named unassigned 11 (located at 2.47 ppm, doublet), show high correlations with the GC-MS data information. Overall, the GC-MS spots (combination of retention time and *m/z* fragments) identified as highly correlated with the described above NMR data were the already described 5-HMF, acetic acid, furfural, diethyl succinate, maltoxazine and unknown 1 (all identified before as varying significantly with aging in Table 5.4) and the newly detected ethyl nicotinate (*m/z* 123 and 156, at 4.62 mins), methyl 2-furoate (*m/z* 95 and 126, at 5.44 mins), γ-nonalactone (*m/z* 49 and 85, at 6.29 mins), a methyl octanoate (tentative assignment, *m/z* 126, at 3.98 mins), a possible glutamic acid product (*m/z* 84, at 7.56 mins) and three still assignment compounds: unknown 2 (*m/z* 101 and 144, at 3.30 mins), unknown 3 (*m/z* 68, at 5.76 mins) and unknown 4 (*m/z* 156, at 5.97 mins). Table 5.5 summarizes the identified correlations and the corresponding compound assignments. Some examples of the NMR/GC-MS 2D correlations contour plots are shown in Figure 5.20 for a) 5-HMF, b) succinic acid, c) branched H1 α (1→4), d) linear H1 α (1→4), e) unassigned 6, and f) unassigned 11. On these correlation plots, the pairs of variables showing a high positive correlation (> 0.8) are represented in red,

whereas the anticorrelated variables (< -0.8) show a blue color. The identified high correlations between NMR and GC-MS variables may be the effect of the presence of the same metabolite, from biochemically related metabolites, or from unrelated metabolites that have a similar abundance pattern across the analyzed samples.

Table 5.5. SHY correlation results between NMR and GC-MS data (correlation coefficients cutoff at ± 0.80). Compounds still unassigned are described as *unassigned* in NMR and *unk.* in GC-MS. * tentatively compound identification; [1] resonances related with linear oligomers; [2] resonances related with branched oligomers.

NMR compounds	**GC-MS compounds correlated positively**	**GC-MS compounds correlated negatively**
5-HMF (9.45ppm)	5-HMF/ diethyl succinate/ethyl nicotinate/ furfural/ glutamic product*/ unk. 4	acetic/ maltoxazine/unk.2/ unk. 3
GABA (2.50 ppm)	maltoxazine/ unk. 3	5-HMF/ diethyl succinate/ furfural/ glutamic product*/ methyl 2-furoate/ γ-nonalactone/ unk. 4
Succinic (2.65 ppm)	5-HMF/ furfural	DDMP/ maltoxazine/ methyl octanoate*/ unk. 2/ unk. 3
H1 β red. (4.65 ppm) [1]	acetic/ maltoxazine/ unk.2/ unk. 3	5-HMF/ furfural/ unk. 4
H1 α red. (5.20 ppm) [1]	-	5-HMF
Linear H1 α (1-4) (5.39 ppm) [1]	acetic/ unk. 3	5-HMF/ furfural/ glutamic product*/ unk. 4
H1 α (1-6) (4.94 ppm) [2]	5-HMF/ furfural/ unk. 4	unk. 2/ unk. 3
Branched H1 α (1-4) (5.34 ppm) [2]	5-HMF/ furfural/ glutamic product*/ unk. 4	Acetic
Un. 1 (1.46 ppm)	-	5-HMF/ diethyl succinate/ furfural/ γ-nonalactone/ unk. 4
Un.6 (4.09 ppm)	5-HMF/ furfural/ glutamic product*/ unk. 4	acetic/ maltoxazine/ unk. 2/ unk. 3
Un.11 (2.47 ppm)	acetic/ maltoxazine/ unk. 2/ unk.3	5-HMF/ furfural/ unk. 4

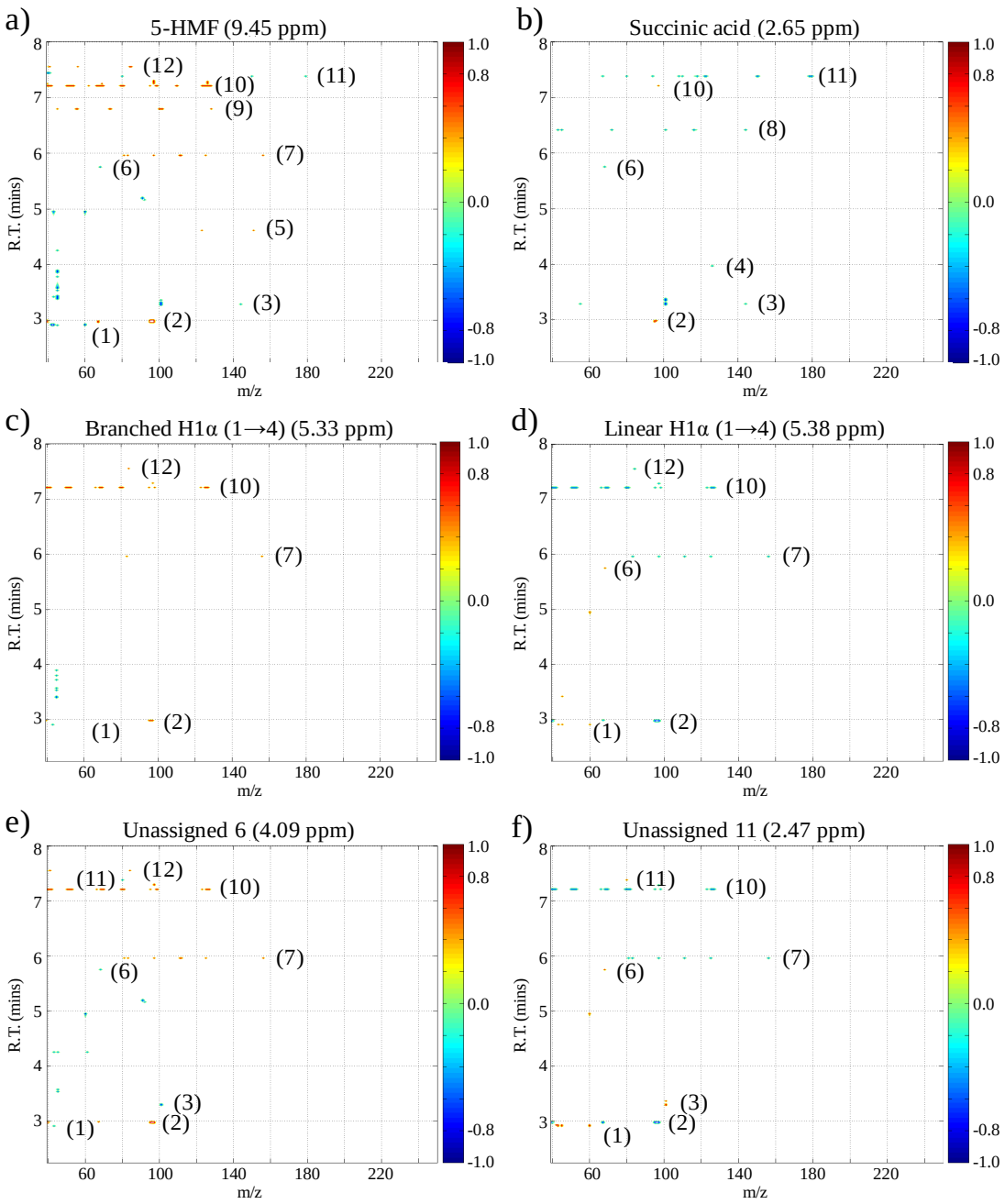

Figure 5.20. 2D contour plots expressing the correlations (cutoff of ± 0.80) between NMR (chemical shifts) and GC-MS data (retention time vs. *m/z* fragments) corresponding to the NMR signals of a) 5-HMF, b) succinic acid, c) branched H1 α (1→4), d) linear H1 α (1→4), e) unassigned compound 6 and f) unassigned compound 11. GC-MS spots identified as correlated are 1) acetic acid, 2) furfural, 3) unknown 2 (*m/z* 101 and 144), 4) possible methyl octanoate (*m/z* 87 and 126), 5) ethyl nicotinate, 6) unknown 3 (*m/z* 68), 7) unknown 4 (*m/z* 125 and 156), 8) DDMP, 9) diethyl succinate, 10) 5-HMF, 11) maltoxazine, and 12) glutamic acid product (*m/z* 41 and 84, possible). In blue: GC-MS spots negatively correlated with the NMR signal; in red: GC-MS spots illustrated as positively correlated with the NMR signal.

As seen before, 5-HMF is clearly the most significant aging marker identified by NMR/PLS-DA, with variation percentage between day 0 and 18 of 150%, and a r > 0.99. Observing the inter-correlation 2D contour plot obtained for its NMR signal (Figure 5.20a), positive correlations (>0.80) were identified with the GC-MS spots corresponding to 5-HMF (high correlation of the same compound detected by different analytical techniques indicating a good response of the inter-correlation methodology), diethyl succinate, ethyl nicotinate, furfural, a possible glutamic acid product and unknown compound 4, whereas acetic acid, maltoxazine and an unknown compound 3 were identified as negatively correlated (<-0.80).These positive correlations between 5-HMF (NMR data) and GC-MS spots 5-HMF, diethyl succinate, ethyl nicotinate and furfural were expected, as these four compounds are well described in literature as increasing with time of storage at an approximately linear rate, being 5-HMF and furfural products of sugars degradation (Vanderhaegen *et al.*, 2003; Vanderhaegen *et al.*, 2006) and diethyl succinate and ethyl nicotinate products of the esterification of acids and alcohols (Vanderhaegen *et al.*, 2003; Vanderhaegen *et al.*, 2006; Saison *et al.*, 2010). Indeed, the compounds 5-HMF, diethyl succinate and furfural were already identified in this study as increasing in an approximately linear manner with the aging process (by GC-MS/PCA). Through peak integration, the ethyl nicotinate increase as a function of aging was confirmed, showing a variation percentage between day 0 and day 18 of aging of 540%, at an approximately linear rate (r = 0.99). Two other compounds were identified as positively correlated, namely a possible glutamic acid product and the unknown 4. By peak integration, an approximately linear increase with aging was observed for both compounds (r values > 0.99), with variation percentages between fresh (day 0) and aged beer at day 18 of 826% and 366% for the unknown 4 and the glutamic acid product, respectively, thus suggesting that these compounds could be also good indicators of the aging process.

The GC-MS spots corresponding to acetic acid, maltoxazine and unknown 3 were identified as highly negatively correlated with 5-HMF (NMR data). Maltoxazine, previously identified as decreasing with aging by GC-MS/PCA, and unknown 3 showed a clear decrease with the aging process, with variation percentages between beer at day 0 and at day 18 of -91% and -78%, respectively. These decreases suggest that these two compounds may be important substrates for Maillard reactions, leading to the formation of heterocyclic compounds (e.g. furfural and 5-HMF). The acetic acid spot, although

identified before by GC-MS/PCA as decreasing with aging, by signal integration no real aging tendency was identified for that compound, thus suggesting that the high correlations identified with this spot (seen in the NMR variables corresponding to 5-HMF, dextrins and unassigned compounds 6 and 11) could be an artefact.

The characteristic NMR resonances of linear (H1 resonances for β red., α red. and α (1-4) in linear oligomers) and branched (H1 resonances for α (1-6) and α (1-4) in branched oligomers) dextrins also showed high correlations with the GC-MS data. Considering the 2D correlation contour plots correspondent to linear dextrins, example of the H1 α (1-4) in linear oligomers 2D contour plot is shown in Figure 5.20d, the GC-MS spots identified as highly correlated correspond to acetic acid, maltoxazine, unknowns 2 and 3 (positively) and to 5-HMF, furfural, glutamic acid product and unknown 4 (negatively). On the other hand, the branched dextrins, example shown in Figure 5.20c for H1 α (1-4) in branched oligomers, show high correlations with the same GC-MS spots, although with opposite signals. The only exception was the GC-MS spot corresponding to maltoxazine, for which no significant correlation with branched dextrins resonances was detected.

The positive correlations between 5-HMF and furfural, well known products of sugars degradation, and branched dextrins (and consequently negative correlations with linear ones) confirm the observations registered through NMR/PLS-DA. In NMR/PLS-DA, a decrease of the average oligomers chain length and an increase of the number of branching points per molecule occur during the aging process was detected, suggesting that oligomers degradation occurs through a mechanism that attacks the shorter linear oligomers to feed Maillard pathways, such as the formation of 5-HMF and furfural. Furthermore, the correlations encountered between the oligomer resonances and the glutamic acid product may also suggest that glutamic acid product formation is dependent on or occurs through Maillard reactions between sugars and glutamic acid (a simplified and general overview of the Maillard reactions is illustrated in Chapter 1, Figure 1.6).

GABA, the only amino acid for which high correlations with GC-MS data were detected, showed a NMR resonance (2.50 ppm) well correlated with 5-HMF, diethyl succinate, furfural, glutamic acid product, methyl 2-furoate, γ-nonalactone and unknown 4 (negatively) and maltoxazine and unknown 3 (positively). The observed correlations between GABA and maltoxazine and unknown 3 (possible Maillard

substrates, positively correlated) and 5-HMF, furfural, glutamic acid product and methyl 2-furoate (Maillard reactions products, negative correlations) suggest GABA as an important substrate for Maillard reactions, thus supporting the suggestion given by 2D NMR correlation analysis. The newly identified compound methyl 2-furoate (*m/z* 95 and 126) is a known product of thermal degradation of sugars (Maga and Katz, 1979) showing an approximately linear increase as a function of aging days, with a variation percentage between fresh beer and beer at day 18 of 235%. GABA also showed negative correlations with some well established markers of beer aging, diethyl succinate (mentioned before) and γ-nonalactone. The γ-nonalactone is a cyclic ester, known to increase in concentration with beer aging, being considered important for the aged beer flavor (coconut, peach, fruity) (Gijs *et al.*, 2002a; Vanderhaegen *et al.*, 2006).

The identified correlations between succinic acid (NMR domain) and GC-MS data also indicate a close relationship between succinic acid and Maillard reactions, as the main identified correlations correspond to the Maillard reaction related-compounds 5-HMF, furfural (positively) and DDMP, maltoxazine and unknowns 2 and 3 (negatively). Furthermore, a negative correlation was detected between succinic acid and possibly methyl octanoate (*m/z* 87 and 126, variation percentage between day 0 and day 18 of 13%), suggesting a decrease of the mentioned ester during storage due to hydrolysis (Vanderhaegen *et al.*, 2007).

Finally, in three still unassigned NMR signals, named 1, 6 and 11, correlations with the GC-MS data were detected. In the 2D contour plots corresponding to unassigned 6 and 11, shown in Figure 5.20e,f respectively, it is interesting to see that they show inter-correlations with the same GC-MS spots, although showing opposite correlations signals. Indeed, the spots corresponding to 5-HMF, furfural, glutamic acid product and unknown 4 show positive correlations (negative in the case of unassigned 11), whereas acetic acid, maltoxazine and unknowns 2 and 3 are negatively correlated with the NMR unassigned compound 6 (positive with unassigned 11). 5-HMF, furfural, glutamic acid product, maltoxazine and unknown 2, 3 and 4 were mentioned before as Maillard reaction related-compounds, suggesting that the unassigned signals 6 and 11 can also be described as possible substrate and product of Maillard reactions, respectively. The unassigned compound 1 is characterized by negative correlations with the well established aging markers 5-HMF, diethyl succinate, furfural and γ-nonalactone

and with the unknown compound 4, suggesting that this compound could be an important marker of fresh beer (as mentioned before by NMR).

Table 5.6 summarizes the compositional changes that have been identified as related with beer aging, by combining the analytical techniques and multivariate analysis, namely NMR/MVA (NMR/PLS-DA, NMR/PCA and 2D correlation analysis), GC-MS/MVA, data fusion by PCT-OPA-PLS-DA and SHY between NMR and GC-MS data. Overall, 28 compounds were identified as varying with aging, with 11 compounds detected by NMR, namely 5-HMF, GABA, proline, higher alcohols, acetic, formic, gallic, pyruvic and succinic acids, and 16 by GC-MS, included 5-HMF, diethyl succinate, furfural, ethyl nicotinate, γ-nonalactone and phenylacetaldehyde (aging markers), methyl 2-furoate, DDMP, maltoxazine, a glutamic acid product, a ketone and methyl octanoate (tentative assignment for these 3 compounds). There are still 10 unassigned relevant compounds (6 related with NMR data and 4 with GC-MS). The ratio linear/branched dextrin was also identified by NMR as varying with beer aging.

The compounds 2,3-butanediol, benzeneacetic acid, 2-phenylacetate, caproic and caprylic acids although identified by GC-MS/MVA methods as varying with the aging process, by signals integration no trend is observed.

Table. 5.6. Summary of the results obtained by NMR and GC-MS data combined with multivariate analysis. ↑ or ↓ refer to compounds showing an increasing or decreasing variation with aging.

NMR data		
Compounds	**δ [a]/ ppm**	**Variation tendency**
Acetic	2.07, s	↑ (< 5%, days 1-7)
Formic	8.23, s	↑ (< 5%, days 1-18)
GABA	2.51, t	↓ (6%, days 0-18)
Gallic	7.09, s	↓ (< 5%, days 2-18)
Higher alcohols [a]	0.86-0.89	↑ (< 5%, days 1-13)
5-HMF	9.45, s	↑ (150 %, days 0-18)
Proline	2.10, m	↓ (< 5%, days 1-18)
Pyruvic	2,40, s	↑ (< 5%, days 0-7)
Succinic	2.65, s	↑ (< 5%, days 0-7)
Unassigned 1	1.46, d	↓ (18%, days 1-13)
Unassigned 3	2.31, d	↑ (< 5%, days 0-10)
Unassigned 5	3.79, t	↑ (< 5%, days 1-13)
Unassigned 6	4.09, d	↑ (13%, days 0-18)
Unassigned 7	6.35, s	↓ (39%, days 0-18)
Unassigned 11	2.47, d	↓ (< 5%, days 1-18)
Av. no. branching points/ molecule	-	↑ (16%, days 0-18)
Av. no. glucose units/ molecule	-	↑ (5%, days 0-7)

GC-MS data			
	R.T./mins	**Main m/z fragments**	**% variation (day 0 -18)**
Acetic acid [b]	2.9	43	-
2,3-butanediol [b]	3.3	45	-
Benzeneacetic acid	7.5	91/ 136	-
Caproic acid [b]	4.8	60	-
Caprylic acid [b]	5.6	60	-
DDMP	6.4	43/ 101/ 144	↓ (42%)
Diethyl succinate	6.8	101	↑ (353%)
Ethyl nicotinate	4.6	123/ 151	↑ (540%)
Furfural	3.0	67/ 95	↑ (946%)
Glutamic product [c]	7.6	41/ 84	↑ (366%)
5-HMF	7.2	69/ 97/ 126	↑ (218%)
Ketone [c]	5.9	43/ 114	↓ (8%)
Maltoxazine	7.3	150/ 179	↓ (91%)
Methyl 2-furoate	5.4	95/ 126	↑ (235%)
Methyl octanoate [c]	4.0	60/ 74/ 87/ 126	↑ (13%)
γ-nonalactone	6.3	49/85	↑ (227%)
Phenylacetaldehyde	3.9	91	↑ (28%)
2-phenylethyl acetate [b]	4.6	104	-
Unknown 1	4.3	43	↓ (11%)
Unknown 2	3.3	101/ 144	↓ (87%)
Unknown 3	5.8	49/ 68/ 83	↓ (78%)
Unknown 4	6.0	97/111/ 125/ 156	↑ (826%)

[a] ppm range comprising overlapped signals arising from isobutanol, isopentanol and propanol. [b] Compounds for which no clear aging trend was identified by signal integration. [c] Tentative assignment.

5.6.3 Analysis of sensory/NMR and sensory/GC-MS data sets by PLS1 regression

Exploratory sensory analysis was performed by four tasters recruited with basis on their good sensory ability and experience on the identification of stale off-flavors in aged beer. Figure 5.21 illustrates the boxplot of the sensory response as a function of the days of aging. It may be seen that day 7 was the day at which the threshold limit for rejecting the beer (identified by a red line) was crossed. Although the analysis was only exploratory (as the number of tasters was limited to only four tasters), day 5 was assumed as the "border day" for organoleptically acceptable beer. To avoid the influence of over-aged beer samples, only the beer samples ranging from day 0 to day 10 of aging were used in the built up of PLS1 regression models.

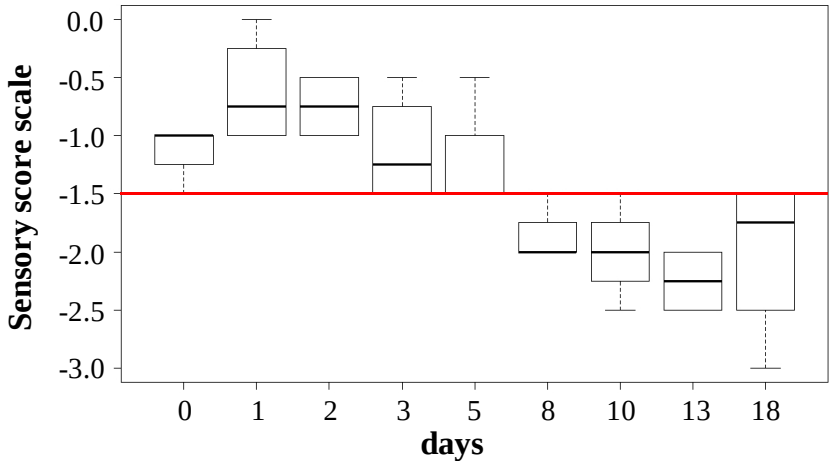

Figure 5.21. Boxplot of the sensory response as a function of the days of aging. Red line expresses the threshold limit for rejecting the beer.

PLS1 regression models, using as reference the sensory response, were built for NMR, using each spectral region: aliphatic (0.8-3.1 ppm), anomeric (4.5-5.8 ppm) and aromatic (5.8-10.0 ppm) regions (the latter shown in Figure 5.22a, as it showed better model performance, when compared with the other spectral regions), and GC-MS data sets (illustrated in Figure 5.22b). The obtained results are summarized in Table 5.7.

All the PLS1 regression models built were found to correspond to three latent variables (LVs). The obtained models are characterized by accumulated R^2Y ranging from 79.6 (for the NMR aliphatic region) to 95.6% (for the NMR aromatic region) and by exhibiting good correlation coefficients ($r \geq 0.89$) with the highest value ($r = 0.98$) being obtained, once again, for the aromatic region, meaning that the aromatic region

spectral information seems to be more related with the changes observed by sensory analysis. The prediction power of these PLS1 regression models was relatively low, as expressed by the high values (> 20%) of the root mean square error of prediction (RMSEP), possibly reflecting the use of a reduced number of samples and tasters for sensory analysis.

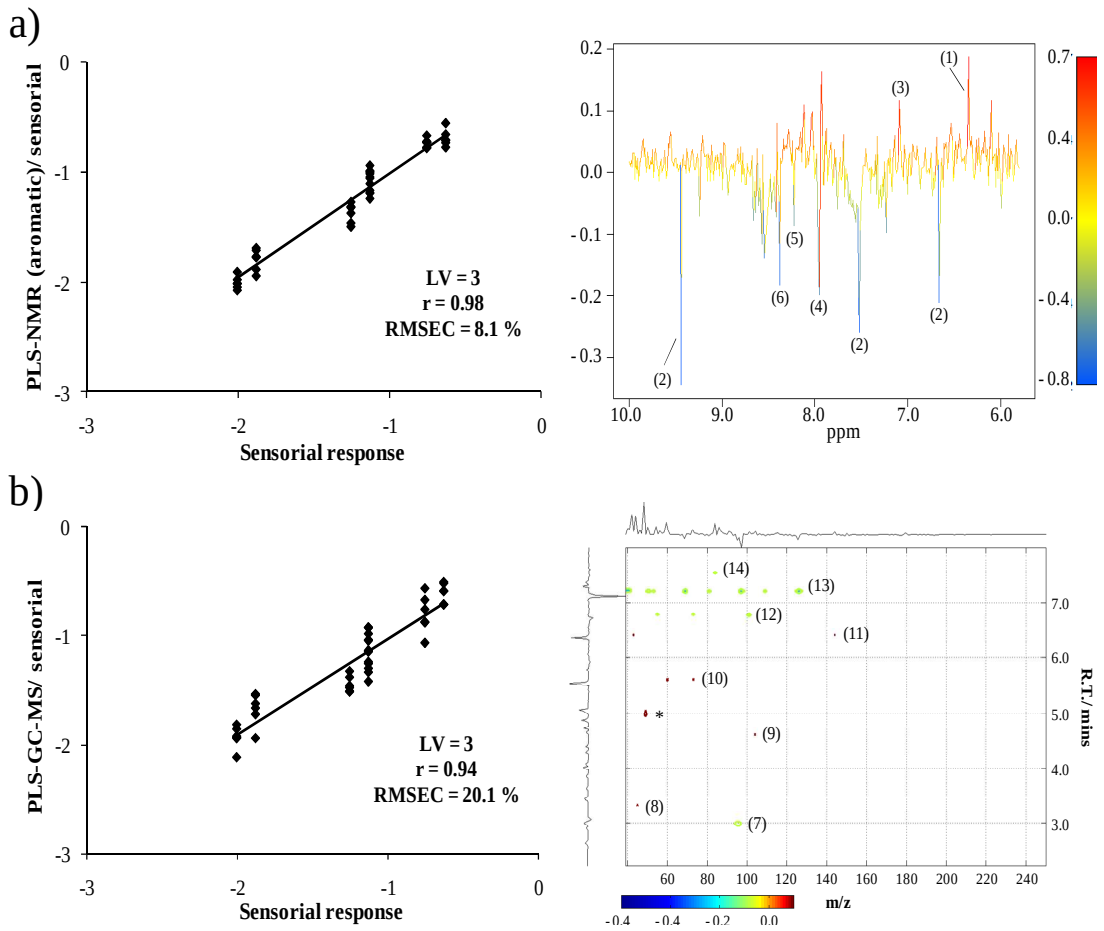

Figure 5.22. PLS1 regression models using the sensory response as reference (left) and corresponding loadings weights (w1) color plots (right) obtained for a) NMR aromatic (5.8-10.0 ppm) spectral region data and b) rapid GC-MS data, as a function of days of aging (day 0-day 10). Identified compounds refer to (1) unassigned 7, (2) 5-HMF (peaks at 6.67, 7.52 and 9.45 ppm), (3) gallic acid, (4) unassigned 8, (5) formic acid, (6) unassigned 9 (NMR domain) (unassigned compounds defined previously in Figure 5.3), (7) furfural, (8) 2,3-butanediol, (9) 2-phenylethyl acetate, (10) caproic acid, (11) DDMP, (12) diethyl succinate, (13) 5-HMF and (14) glutamic acid product (GC-MS data). Color plot reflects the correlation with sensory response: red and blue expresses, respectively, positive and negative correlations with the fresher beer. *: Artefact related with the solvent (dichloromethane).

To interpret the variability contained in each PLS1 regression model, the corresponding loadings weights (**w1**) profiles, examples of aromatic NMR spectral region and GC-MS are shown in Figure 5.22 (right), were analyzed. The profiles are colored as a function of the correlation between each variable and the **y** vector data (sensory response), with positive correlations (signals decreasing with aging) being shown in red, whereas negatively correlated (signals increasing with aging) are shown in blue and green. The variables closely related with the sensory data are summarized in Table 5.7.

Considering the loadings weights (**w1**) profiles obtained, it was noted that the signals showing higher correlations with the sensory response correspond to 5-HMF, GABA, proline, ratio linear/ branched dextrins, acetic, citric, formic, gallic, pyruvic acid and hydrate form and succinic acids and unassigned compounds 1, 7, 8 and 9 for the PLS1-NMR models, and to 2,3-butanediol, 5-HMF, caproic acid, DDMP, diethyl succinate, furfural, glutamic acid product and 2-phenylethyl acetate spots for the PLS1-GC-MS (Figure 5.22, right) (all those compounds identified before as varying with the aging process, by NMR/MVA and GC-MS/MVA).

Table 5.7. PLS1 regression models results obtained for sensory/NMR and sensory/GC-MS data.

	LV	**R^2Y (%)**	**RMSECV (%)**	**RMSEP (%)**	**R**	**Signals related with sensory response**
NMR (aliphatic)	3	83.8	15.5	27.2	0.92	Acetic, citric, GABA, proline, pyruvic and hydrate form, succinic and unassigned 1
NMR (anomeric)	3	79.6	17.4	22.9	0.89	Branched/linear dextrins
NMR (aromatic)	3	95.6	8.1	24.2	0.98	Formic, gallic, 5-HMF and unassigned 7, 8 and 9
GC-MS	3	87.5	13.6	20.1	0.94	2,3-butanediol, caproic, DDMP, diethyl succinate, 5-HMF, furfural, 2-phenylethyl acetate and glutamic acid product

In order to improve both selectivity as well as the predictive ability of the built PLS1 regression models, future studies should be performed using not only a higher number of beer samples but also of beer tasters (at least 7). Overall, the built up of PLS1

regression models enabled achieving promising results in the application of this methodology in order to achieve deeper understanding of beer aging process as well as to establish a bridge between the organoleptic characteristics of aged beer and analytical techniques.

5.6.4 Conclusions of data correlation analysis between NMR and GC-MS, sensory/NMR and sensory/GC-MS data sets

PCT-OPA-PLS-DA data fusion and SHY methodologies were applied to NMR and GC-MS data sets in order to combine the highest sensitivity of GC-MS, allowing the identification of several aging markers, with the NMR ability to detect wide range of families of compounds in a single run, showing higher reproducibility.

The results obtained by PCT-OPA-PLS-DA applied to NMR and GC-MS data enabled the identification of a clear aging trend, very similar to the one obtained by GC-MS/MVA (section 5.4), thus indicating the highest weight of the GC-MS data on the resulting outer-product matrix. Although affected by that weight difference, OPA methodology was able to identify the same compounds varying with aging process than when MVA is applied directly to each data set, namely: acetic, formic, gallic, pyruvic and succinic acids, GABA, proline, higher alcohols, 5-HMF, anomeric protons H1 (α and β red., α (1-6) and α (1-4)) and five still unassigned compounds (NMR/ MVA); and acetic, 2,3-butanediol, caproic and caprylic acids, DDMP, diethyl succinate, furfural, 5-HMF, 2-phenylethyl acetate and unknown compounds 1, 2 and 3 (GC-MS/MVA). Furthermore, benzeneacetic acid and a possible ketone (tentative assignment), from GC-MS domain, were also identified as varying with the aging process.

The application of SHY methodology revealed significant inter-compound dependences between the studied NMR and GC-MS domains. For example, the variations of GABA, succinic acid and the ratio linear/branched dextrins (seen in NMR domain) were highly correlated with the ones of known aging markers (GC-MS domain) 5-HMF, furfural, 2-methyl-furoate, a glutamic product (Maillard reaction products), diethyl succinate and γ-nonalactone, as well as of maltoxazine (substrate of Maillard reactions) and an unknown compound 3, indicating a major influence of Maillard reactions during aging. Taking into account the observed correlations, besides the confirmation of the relevance of several known aging markers and their

dependencies, a interesting relation between Maillard and esterification reaction products was identified, indicating a common or a similar formation rate between Maillard and the esterification reactions. Furthermore, SHY approach enabled a more complete description of the compositional changes occurring during the aging process, by identifying as correlated with the aging process the known aging markers ethyl nicotinate, methyl-2-furoate and γ-nonalactone, as well as five other compounds, methyl octanoate (assignment not confirmed), a possible glutamic acid product and 3 still unassigned compounds.

Overall, the combination of NMR and GC-MS data sets by different MVA strategies enabled the identification of 29 compounds, the ratio linear/branched dextrins and 10 still unassigned compounds as varying with the aging process. Indeed, in GC-MS domain, both well established aging markers, such as 5-HMF, diethyl succinate, furfural, ethyl nicotinate, γ-nonalactone and phenylacetaldehyde, as well as other relevant compounds, as DDMP, maltoxazine, methyl 2-furoate and four unassigned compounds, showing significantly high variation percentage between fresh beer (day 0) and beer at day 18 of aging, varying between 30% (for phenylacetaldehyde) and 950% (for furfural). The compositional variations identified in NMR are of lower magnitude, with only 5-HMF showing high percentage variation (approximately 150%), however showed great potential to assess the aging process as it allows the detection of a great number of family compounds in a single run, with a simple and fast preparation step and a high reproducibility. In the future, exhaustive characterization studies have to be performed in order to identify the relevant but still unassigned compounds that were described above. These compounds not only have the potential, in some cases, to be aging markers, but can also give a more detailed picture of the compositional steps and chemical pathways that are varying during the aging process.

Beer aging process was also followed by sensory analysis, which enabled to define day 5 as the “border day” to organoleptically acceptable beer. Taking into account that information, exploratory PLS1 regression models were built using the sensory response as reference. The observed results showed good relationships between analytical and sensory data ($r > 0.89$), especially for the aromatic region of the NMR spectra and for GC-MS data, with correlations of 0.98 and 0.94, respectively. This may suggest the use of NMR and GC-MS data to predict the degree of beer sensory deterioration. However, for all PLS1 regression models built high values of RMSEP, >

20%, were observed indicating a relatively low prediction power, reflecting the reduced number of samples and of tasters in the study.

6. FINAL CONCLUSIONS AND FUTURE PERSPECTIVES

6. Final Conclusions and Future Perspectives

High resolution NMR spectroscopy has proven to be a powerful method for the identification and characterization of a variety of components by direct analysis of beer. That potential was explored in a large part of this thesis for assessing beer quality and stability. Indeed, the 1D and 2D NMR spectra, as well as LC-NMR and LC-MS data, obtained by direct analysis of beer revealed a great richness and complexity of information about beer composition. In this way and, 40 compounds were assigned by NMR methods, present in a wide range of concentrations, including alcohols (ethanol, glycerol, isobutanol, isopentanol, 2-phenylethanol, propanol and tyrosol), amino acids (three letter code: Ala, Arg, Glu, Gln, His, Ile, Leu, Lys, Phe, Pro, Thr, Try, Val and GABA), nucleosides (adenosine, cytidine, inosine and uridine), organic acids (acetic, citric, formic, gallic, lactic, malic, pyruvic and succinic acids), sugars (maltose and trehalose) and other relevant compounds as acetaldehyde and 5-HMF as well as dextrins resonances. However, about 20 spin systems still remain unknown, calling for improved assignment. Although no further compound assignment was achieved when compared to literature, the fact that new samples preparation were here employed (pH adjustment to ~1.90) made this detailed characterization extremely important for the metabonomic studies performed in this thesis.

MVA was applied to ^{1}H NMR data (NMR/MVA) in order to evaluate beer variability between beers from the same brand, produced nationally but differing in brewing sites (sites **A**, **B**, and **C**) and dates (**a**, **b** and **c**), aiming at the short-term monitoring of beers. A previous report has addressed this issue, however brewing sites were located at different countries (Almeida *et al.*, 2006), and where a separation between beers differing in brewing sites was observed based on differences in brewing steps (malting and mashing steps), besides the fact that water fact origin is distinct in different countries. In this thesis, the beers studied were produced nationally, thus ensuring a better control and homogeneity of the starting materials (e.g. water used) as well as of the brewing protocols, leading to an overall finer control of the process and of the final composition of beer. The corresponding NMR/MVA study enabled to detect differences on the production reproducibility performance of brewing sites, with site **A** beers showing great compositional reproducibility, whereas beers from site **B** demonstrated extensive variability within dates of production. Furthermore,

NMR/MVA revealed compositional differences between the analyzed beers, indicating differences between brewing sites and/or dates in several production steps, such as the fermentation extent (measured by the relationship between amino acids content and amounts of higher alcohols and organic acids and by the levels of acetaldehyde), mashing step (due to nucleosides content), efficiency of maturation (levels of acetaldehydes) and mashing process (ratio branched/linear dextrins), as well as a potential microbiological contamination (lactic acid content).

The sensitivity shown by the NMR/MVA method for the identification of finer compositional variations within brewing sites has revealed promising for a future routine application in an industrial environment, as an automated high-throughput method to control brewing process. Such a methodology has been applied to the fruit juices sector with great success (Spraul *et al.*, 2009).

The quantification of the main organic acids (i.e. acetic, citric, lactic, malic, pruvic and succinic acids) present in beer by NMR methods was assessed, aiming the development of quantitative high throughput methods for potential routine applications in the brewing process control, as organic acids are good indicators of the fermentation performance. Although inherently quantitative, NMR of complex mixtures such as beer, faces problems due to i) signal overlap, and ii) the use of quantitative conditions (use of inter-scan times at least 5 x highest T_1), thus increasing drastically the time of analysis (40 minutes per spectra vs. 15 minutes per spectra in semi-quantitative conditions). Partial least squares (PLS) regression method has been presented as a suitable method for rapid quantification, allowing a routine application to large sample populations. This procedure, already used in beer to quantify amino acids and organic acids (Nord *et al.*, 2004; Lachenmeier *et al.*, 2005), enables the use of shorter NMR runs through semi-quantitative spectra (15 minutes per sample), circumventing the difficulties associated with long acquisition times and errors associated to signal integration.

In this thesis, different quantitative NMR methodologies were developed and compared, by direct integration methods (*vs.* TSP or ERETIC method) and by PLS regression models. In this way, PLS1-NMR regression models were built using different reference analytical methods: capillary electrophoresis (CE) direct detection and indirect detection, and enzymatic essays, the latter two showing best performance. By comparison of the predictive power of those best performance PLS1-NMR models with

the results obtained through NMR integration methods, either using an internal reference or the ERETIC method, a good agreement was observed, except for malic and pyruvic, for which integration overestimates concentrations (probably due to additional underlying resonances), and for citric acid, for which an apparent overestimation by PLS1-NMR is observed. Overall, the less time consuming method of PLS1-NMR was found to be suitable for organic acid quantification in beer, as long as the best combination of organic acid and particular PLS1-NMR model (differing in the reference method used for NMR calibration) is employed. Although opening an interesting possibility for future routine application in the quantification of organic acids, for the sake of statistical robustness of the quantitative method, larger batches of samples must be considered in the construction of those models.

Finally, metabonomic studies were performed to better understand the forced aging beer process. In this way, MVA methods were applied to several analytical techniques data sets: NMR spectroscopy, GC-MS and MIR spectroscopy, to assess the major compositional changes occurring in beer during the aging process. The use of NMR spectroscopy and GC-MS, which work at different quantitative domains with GC-MS showing higher sensitivity, was based on their great ability to provide compound structural information, enabling achieving new insights into beer aging chemistry and unveiling possible new compounds and potential markers of the process. GC-MS has been extensively used in beer aging studies, with several aging markers being detected by this technique, such as phenylacetaldehyde, methional, 5-HMF, furfural and diethyl succinate. On the other hand, the detailed accompanying of beer aging was here performed, for the first time, by ^{1}H NMR spectroscopy of direct beer analysis. MIR spectroscopy was also used in an attempt to achieve a fast methodology for assessing the degree of beer aging, with future potential application in the industry. Furthermore, the aging process was also followed by sensory analysis, in order to establish a bridge between the organoleptic (sensory response) and the compositional changes (identified by the analytical techniques) occurring during the aging process, in order to achieve a deeper understanding of aging process and how the compositional changes affect the organoleptic properties of beer. Taking into account the literature, and after optimization of the experimental design for beer forced aging, it was decided to use as forced aging conditions: 18 days, at 45 (±1) ºC.

By NMR/MVA, a clear aging trend was observed, with aged beers being distinguished from fresh ones with basis in higher amounts of 5-HMF, acids and the decreases of GABA, gallic acid and the ratio linear/branched dextrins. The latter was, to our knowledge, for the first time reported here, suggesting that dextrin degradation occurs during aging in order to feed the Maillard pathways, and preferentially involving the shorter, linear oligomers. Furthermore, 2D correlation analysis was also performed to assess relationships within NMR data. This approach, applied here for the first time in beer, enabled the mentioned relevant compound changes to be confirmed in a rapid manner and revealed interesting inter-compound dependences, such as between succinic acids, pyruvic and one unassigned signal with GABA, suggesting GABA as a important substrate of Maillard reactions.

In GC-MS analysis, a fast GC-MS/MVA methodology was developed, enabling the reduction of chromatographic separation time, without losing relevant spectral information and the semi-quantitative ability to characterize the detected compounds. To assess that no or minimal information loss occurs, both conventional GC-MS, i.e. working at optimum chromatographic resolution, and rapid GC-MS runs were acquired and compared, showing equivalent compositional information and, thus indicating no significant loss of information. The subsequent GC-MS/MVA methodology enabled the detection of a clear degree of aging of beer, with the identification of known markers 5-HMF, furfural, diethyl succinate and phenylacetaldehyde as increasing with the process. Furthermore, for the first time, DDMP and maltoxazine were identified as decreasing with aging, suggesting that they could take part in the Maillard reactions.

Regarding MIR/MVA studies, a first approach was the use of direct MIR analysis, i.e. simple acquisition of MIR spectra. However, the corresponding spectrum of beer was dominated by water signals, although signals corresponding to branched and linear dextrins and ethanol could also be detected. The performed direct MIR/MVA models were not able to distinguish an aging degree trend, although a separation within days of aging was observed. To overcome the water issue, an alternative methodology was attempted. This methodology, termed droplet evaporation, consisted on the evaporation of a beer droplet and subsequent spectral acquisition, allowing the decrease of the water bands and subsequent enhancement of the remaining signals particularly branched dextrins to be achieved. As in the case of direct MIR, the droplet evaporation MIR/MVA models were only able to distinguish beers between studied days, however

no aging trend was detected. Although the droplet evaporation methodology seems promising to overcome the water influence limitations, low reproducibility of the evaporation process was detected and, therefore, developments and/or optimization are required if application of this methodology is to be envisaged.

Data correlation between NMR and GC-MS data sets has been performed by applying the MVA methods of principal component transform-outer product analysis-partials least squares-discriminant analysis (PCT-OPA-PLS-DA) data fusion and statistical heterospectroscopy (SHY). In PCT-OPA-PLS-DA methodology, a clear aging trend was identified, very similar to the one observed for GC-MS/MVA, indicating different weightings between of GC-MS (significantly higher) and NMR data sets as a function of aging. Although affected by that weightings difference, PCT-OPA-PLS-DA was able to identify the same compounds varying with aging process than when MVA is applied directly to each data set, as well as benzene acetic acid and a possible ketone (tentative assignment).

The SHY methodology was applied to identify highly correlated pairs of variables between NMR and GC-MS domains. Using as a correlation value threshold ± 0.80, several significant inter-compound dependences were identified. For example, the ongoing Maillard reactions were well characterized through correlations between 5-HMF, furfural, GABA, maltoxazine, methyl 2-furoate, γ-nonalactone, succinic acid and the ratio linear/branched dextrins. Inter-dependency between Maillard reaction compounds (5-HMF) and esterification reaction products (diethyl succinate and ethyl nicotinate) was also identified, indicating a common or a similar formation rate between Maillard and the esterification reactions. Furthermore, SHY methodology enabled a more complete description of the compositional changes occurring during the aging process by identifying ethyl nicotinate, methyl 2-furoate and γ-nonalactone (known aging markers), as well as five other compounds: methyl octanoate (assignment not confirmed), a possible glutamic acid product and 3 still unassigned compounds.

Overall, data correlation of NMR and GC-MS data sets enabled the identification of 29 compounds varying with the aging process, including known markers 5-HMF, diethyl succinate, furfural, ethyl nicotinate, γ-nonalactone and phenylacetaldehyde, as well as methyl 2-furoate, DDMP and maltoxazine (the latter two to our knowledge, for the first time identify as relevant in aging process) (GC-MS domain) and GABA, proline, higher alcohols, acetic, formic, gallic, pyruvic and

succinic acids, as well as the ratio of linear/branched dextrins (NMR domain). Furthermore, 10 unassigned compounds were identified as varying with aging, calling for improved assignment.

Data correlation between sensory response and NMR and GC-MS data has also been assessed. Sensory analysis was performed by four tasters specialized in the identification of stale off-flavors in aged beer, which defined day as the "border day" to organoleptically acceptable beer. With basis in that information, PLS1 regression models were built for each NMR spectral region: aliphatic, sugar and aromatic regions, as well as for the GC-MS data, using as reference the sensory response. The results obtained showed good relationships between analytical data response and sensory response ($r > 0.89$), especially for the aromatic region of the NMR spectra and for GC-MS data, with correlations of 0.98 and 0.94, respectively, suggesting that the information contained in those data sets are more well-related with the changes observed by sensory analysis. However, the prediction power of all PLS1 regression models built was relatively low, as expressed by high values of RMSEP (> 20%), reflecting the use of a reduced number of samples and tasters for the study.

In summary, the possibility of direct NMR analysis of beer, in a high throughput manner, opens interesting possibilities for future routine applications in an industrial environment. In fact, the use of MVA tools for the interpretation of NMR data, the so-called metabonomics approach, enables large batches of samples to be considered, in order to develop qualitative and/or quantitative models, statistically robust and having predictive power for subsequent routine analysis of new samples. Furthermore, the in tandem use of NMR and MS data has proven in this work to be a promising strategy to circumvent the inherently low sensitivity of the NMR, so that future metabonomic-based developments will benefit from the parallel use of the two tools.

In the context of beer analysis and quality control, in this thesis, compositional deviations related to several quality parameters (including production site and date and beer aging) have been identified and their relationships assessed with the objective ofestablish the basis for the definition of improved beer quality markers.

At shorter term, several possible avenues may be envisaged for future developments in the area of beer analysis: i) further insight into the characteristics of beer raw materials and their impact on the final product, either for quality control or

informed development of new products, ii) on-line assessment of the compositional changes taking place through the different stages of brewing (e.g. mashing, fermentation, and maturation steps), for improved process control and process understanding, and iii) development of valid, robust correlation models between NMR data and determinant beer sensory quality parameters, towards the objective understanding and the evaluating of sensory properties of the final product.

REFERENCES

Aberg, K.M.; Alm, E.; Torgrip, J.O. The correspondence problem for metabonomics datasets. *Analytical and Bioanalytical Chemistry*, **2009**, 394, 151-162.

Agabriel, D., Cornu, A., Journal, C., Sibra, C., Grolier, P., Martin, B. Tanker milk variability according to farm feeding practices: vitamins A and E, carotenoids, color and terpenoids. *Journal of Dairy Science*, **2007**, 90, 4884-4896.

Akoka, S.; Trierweiler, M. Improvement of the ERETIC method by digital synthesis of the signal and addition of a broadband antenna inside the NMR probe. *Instrumentation Science and Technology*, **2002**, 30, 21-29.

Akoka, S.; Barantin, L.; Trierweiler, M. Concentration measurement by proton NMR using the ERETIC method. *Analytical Chemistry*, **1999**, 71, 2554-2557.

Albert, K. *On-Line LC-NMR and Related Techniques*; John Wiley & Sons: Chichester, 2002.

Almeida, C.; Duarte, I.F.; Barros, A.; Rodrigues, J.; Spraul, M.; Gil, A.M. Composition of beer by ^{1}H NMR spectroscopy: effects of brewing site and date of production. *Journal of Agricultural and Food Chemistry*, **2006**, 54, 700-706.

Altria, K. Applications of standard methods in capillary electrophoresis for drug analysis. In *Handbook of Analytical Separations*, Ed. Valko, K.; Elsevier Science B.V., 2000, 87-105.

Andersen, M.L.; Outtrup, H.; Skibsted, L.H. Potential antioxidants in beer assessed by ESR spin trapping. *Journal of Agricultural and Food Chemistry*, **2000**, *48*, 3106-3111.

Aue, W.P.; Karhan, J.; Ernst R.R. Homonuclear broad band decoupling and two-dimensional J-resolved NMR spectroscopy. *Journal of Chemical Physics*, **1976**, 64, 4226-4227.

Bamforth, C.W. *Beer: Tap into the Art and Science of Brewing*, Oxford University Press: New York, 2nd Edition, 2003.

Bamforth, C.W. *Beer: A Quality Perspective*; Academic Press: San Diego, 2009.

Barantin, L.; Le Pape, A.; Akoka, S. A new method for absolute quantitation MRS metabolites. *Magnetic Resonance in Medicine*, **1997**, 38, 179-182.

Barros, A.S.; Safar, M.; Devaux, M.F.; Robert, P.; Bertrand, D.; Rutledge, D.N. Relations between mid-infrared and near-infrared spectra detected by analysis of variance of an intervariable data matrix. *Applied Spectroscopy*, **1997**, 51, 1384-1393.

Barros, A.S.; Rutledge, D.N. Principal component transform-partial least squares: a novel method to accelerate cross-validation in PLS regression. *Chemometrics and Intelligent Laboratory Systems*, **2004**, 73, 245-255.

Barros, A.S.; Pinto, R.; Delgadillo, I.; Rutledge, D.N. Segmented principal component transform-partial least squares regression. *Chemometrics and Intelligence Systems*, **2007**, 89, 59-68.

Barros, A.S.; Pinto, R.; Bouveresse, D.J.R.; Rutledge, D.N. Principal component transform – Outer product analysis in the PCA context. *Chemometrics and Intelligent Laboratory Systems*, **2008**, 93, 43-48.

Barton II, F.E.; Himmelsbach, D.S.; Duckworth, J.H.; Smith, M.J. Two-dimensional vibration spectroscopy: correlation of mid- and near-infrared regions. *Applied Spectroscopy*, **1992**, 46, 420-429.

Becker, E.D. *High Resolution NMR: Theory and Chemical Applications*; Academic Press: SanDiego, 3nd Edition, 2000.

Belke, C.J.; Irwin, A.J. Determination of organic acids in beer after extraction with an anion-exchange resin. *Journal of American Society of Brewing Chemists*, **1992**, 50, 26-29.

Berregi, I.; Santos, J.I.; del Campo, G.; Miranda, J.I.; Aizpurua, J.M. Quantitation determination of chlorogenic acid in cider apple juices by ^{1}H NMR spectrometry. *Analytica Chimica Acta*, **2003**, 486, 269-274.

Black, J.A.; Cunningham, G.; Robson, E. *The literature of ancient Sumer*; Oxford University Press: Oxford, 2004.

Bobalova, J.; Petry-Podgorska, I.; Lastovickova, M.; Chmelik, J. Monitoring of malting process by characterization of glycation of barley protein Z. *European Food Research and Technology*, **2010**, 230, 665-673.

Bock, K.; Duus, J.O.; Norman, B.; Pedersen, S. Assignment of structures to oligosaccharides produced by enzymic degradation of a beta-D-glucan from barley by ^{1}H- and ^{13}C NMR spectroscopy. *Carbohydrate Research*, **1991**, 211, 219-233.

Bodenhausen, G.; Ruben, D.J. Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chemical Physics Letters*, **1980**, 69, 185-188.

Boulton, C.; Quain, D. *Brewing Yeasts and Fermentation*; Blackwell: Oxford, 2006.

Branyik, T.; Vicente, A.A.; Dostalek, P.; Teixeira, J.A. A review of flavour formation in continuous beer fermentations. *Journal of the Institute of Brewing*, **2008**, 114, 3-13.

Braunschweiler, L.; Bodenhausen, G.; Ernst, R.R. Coherence transfer by isotropic mixing: application to proton correlation spectroscopy. *Journal of Magnetic Resonance*, **1983**, 53, 521-528.

Bravo, A.; Sanchez, B.; Scherer, E.; Herrera, J.; Rangel-Aldao, R. α-dicarbonylic compounds as indicators and precursors of flavor deterioration during beer aging. *Master Brewers Association of the Americas Technical Quarterly*, **2002**, 39, 13-23.

Brezova, V.; Polovka, M.; Stasko, A. The influence of additives on beer stability investigated by EPR spectroscopy. *Spectrochimica Acta Part A*, **2002**, 58, 1279-1291.

Brescia, M.A.; Kosir, I.J.; Caldarola, V.; Kidric, J.; Sacco, A. Chemometric classification of Apulian and Slovenian wines using H-1 NMR and ICP-OES together with HPICE data. *Journal of Agricultural and Food Chemistry*, **2003**, 51, 21–26.

Briggs, D.E.; Hough, J.S.; Stevens, R.; Young, T.W. *Malting and Brewing Science: Vol. 1 Malt and sweet wort*. Kluwer Academic: Suffolk, 1981.

Briggs, D.E.; Boulton, C.A.; Brookes, P.A.; Stevens, R. *Brewing: science and practice*. Woodhead Publishing Ltd.: Cambridge, 2004.

Broberg, A.; Thomsen, K.K.; Duus, J.O. Application of nano-probe NMR for structure determination of low nanomole amounts of arabinoxylan oligosaccharides fractionated by analytical HPAEC-PAD. *Carbohydrate Research*, **2000**, 328, 375-382.

Caligiani, A.; Acquotti, D.; Palla, G.; Bocchi, V. Identification and quantification of the main organic components of vinegars by high resolution 1H NMR spectroscopy. *Analytica Chimica Acta*, **2007**, 585, 110-119.

Callemien, D.; Collin, S. Structure, organoleptic properties, quantification methods, and stability of phenolic compounds in beer – a review. *Food Reviews International*, **2009**, 26, 1-84.

Camara, J.S.; Alves, M.A.; Marques, J.C. Changes in volatile composition of Madeira wines during their oxidative ageing. *Analytica Chimica Acta*, **2006**, 563, 188-197.

Casey, G.P. A journey in brewing science and the ASBC. In *Brewing Chemistry and Technology in the Americas*; Gales, P.W., Ed.; ASBC: St Paul, 2007, 99-231.

Cazor, A.; Deborde, C.; Moing, A.; Rolin, D.; This, H. Sucrose, glucose and fructose extraction in aqueous carrot root extracts prepared at different temperatures by means of direct NMR measurements. *Journal of Agricultural and Food Chemistry*, **2006**, 54, 4681-4686.

Cevallos-Cevallos, J.M.; Reyes-de-Corcuera, J.I.; Etxberria, E.; Danyluk, M.D.; Rodrick, G.E. Metabolomics analysis in food science: a review. *Trends in Food Science and Technology*, **2009**, 20, 557-566.

Charlton, A.J.; Farrigton, W.H.H.; Brereton, P. Application of ^{1}H NMR and multivariate statistics for screening complex mixtures: quality control and authenticity of instant coffee. *Journal of Agricultural and Food Chemistry*, **2002**, 50, 3098-3103.

Chen, J.H.; Sambol, E.B.; DeCarolis, P.; O'Connor, R.; Geha, R.C.; Wu, Y.V.; Singer, S. High-resolution MAS NMR spectroscopy detection of the spin magnetization exchange by cross-relaxation and chemical exchange in intact cell lines and human tissue specimens. *Magnetic Resonance in Medicine*. **2006**, 55, 1245–1256

Chevance, F.; Guyot-Declerck, C.; Dupont, J.; Collin, S. Investigation of the α-damascenone level in fresh and aged commercial beers. *Journal of Agricultural and Food Chemistry*, **2002**, 50, 3818-3821.

Claridge, T.D.W. *High Resolution Techniques in Organic Chemistry*; Elsevier Science: Oxford, 1999.

Cloarec, O.; Dumas, M.E.; Craig, A.; Barton, R.H.; Trygg, J.; Hudson, J.; Blancher, C.; Gauguier, D.; Lindon, J.C.; Holmes, E.; Nicholson, J. Statistical total correlation spectroscopy: an exploratory approach for latent biomarker identification from metabolic 1H NMR data sets. *Analytical Chemistry*, **2005**, 77, 1282-1289.

Consonni, R.; Gatti, A. ^{1}H NMR studies on Italian balsamic and traditional balsamic vinegars. *Journal of Agricultural and Food Chemistry*, **2004**, 52, 3446-2450.

Consonni, R.; Cagliani, L.R.; Benevelli, F.; Spraul, M.; Humpfer, E.; Stocchero, M. NMR and chemometric methods: A powerful combination for characterization of balsamic and traditional balsamic vinegar of Modena. *Analytica Chimica Acta*, **2008**, 611, 31-40.

Cortacero-Ramírez, S.; Segura-Carretero, A.; Castro, M.H.B, Fernandez-Gutierrez, A. Determination of low-molecular-mass organic acids in any type of beer samples by coelectroosmotic capillary electrophoresis. *Journal of Chromatography A*, **2005**, 1064, 115-119.

Craig, A.; Cloarec, O.; Holmes, E.; Nicholson, J.K.; Lindon, J.C. Scaling and normalization effects in NMR spectroscopic metabonomic data sets. *Analytical Chemistry*, **2006**, 78, 2262-2267.

Crockford, D.J.; Holmes, E.; Lindon, J.C.; Plumb, R.S.; Zirah, S.; Bruce, S.J.; Rainville, P.; Stumpf, C.L.; Nicholson, J.K. Statistical heterospectroscopy, an approach to the integrated analysis of NMR and UPLC-MS data sets: application in metabonomic toxicology studies. *Analytical Chemistry*, **2006**, 78, 363-371.

Crockford, D.J.; Maher, A.D.; Ahmadi, K.R.; Barrett, A.; Plumb, R.S.; Wilson, I.D.; Nicholson, J.K. 1H NMR and UPLC-MS statistical heterospectroscopy: characterization of drug metabolites (xenometabolome) in epidemiological studies. *Analytical Chemistry*, **2008**, 80, 6835-6844.

Cuny, M.; Vigneau, E.; Le Gall, G.; Colquhoun, I.; Lees, M.; Rutledge, D.N. Fruit juice authentication by ^{1}H NMR spectroscopy in combination with different chemometrics tools. *Analytical and Bioanalytical Chemistry*, **2008**, 390, 419-427.

D'Imperio, M.; Mannina, L.; Capitani, D.; Bidet, O.; Rossi, E.; Bucarelli, F.M.; Quaglia, G.B.; Segre, A. NMR and statistical study of olive oils from Lazio: a geographical, ecological and agronomic characterization. *Food Chemistry*, **2007**, 105, 1256-1267.

Daems, V.; Delvaux, F. Multivariate analysis of descriptive sensory data on 40 commercial beers. *Food Quality and Preference*, **1997**, 8, 373–380.

Dalgliesh, C. E. Flavour stability. *Proceedings of the European Brewery Convention Congress*, **1977**, 623–659.

Davidek, T.; Robert, F.; Devaud, S.; Vera, F.A.; Blank, I. Sugar fragmentation in the Maillard reaction cascade: formation of short-chain carboxylic acids by a new oxidative α-dicarbonyl cleavage pathway. *Journal of Agricultural and Food Chemistry*, **2006**, 54, 6677-6684.

De Graaf, R.A. *In Vivo NMR Spectroscopy – Principles and Techniques*, 2nd Edition, John Wiley & Sons: Chichester, 2007.

De Keukeleire, D. Fundamentals of beer and hop chemistry. *Química Nova*, **2000**, 23, 108-112.

De Schutter, D.P.; Saison, D.; Derdelinckx, G.; Delvaux, F.R. The Chemistry of Aging Beer. In *Beer in Health and Disease Prevention*; Preedy, V.R., Ed.; Elsevier: Oxford, 2008, vol. 1, 375-388

Deborde, C.; Maucourt, M.; Baldet, P.; Bernillon, S.; Biais, B.; Talon, G. ; Ferrand, C.; Jacob, D.; Ferry-Dumazet, H.; de Daruvar, A.; Rolin, D.; Moing, A. Proton NMR quantitative profiling for quality assessment of greenhouse-grown tomato fruit. *Metabolomics*, **2009**, 5, 183-198.

del Campo, G.; Berregi, I.; Caracena, R.; Santos, J.I. Quantitative analysis of malic and citric acids in fruit juices using proton nuclear magnetic resonance spectroscopy. *Analytica Chimica Acta*, **2006**, 556, 462-468.

Depraetere, S.A.; Delvaux, F.; De Schutter, D.; Williams, I.S.; Winderickx, J.; Delvaux, F.R. The influence of wort aeration and yeast preoxygenation on beer staling processes. *Food Chemistry*, **2008**, 107, 242-249.

Dervilly, G.; Leclercq, C.; Zimmermann, D.; Roue, C.; Thibault, J.F.; Saulnier, L. Isolation and characterization of high molar mass water-soluble arabinoxylans from barley and barley malt. *Carbohydrate Polymers*, **2002**, 47, 143-149.

Diaz-Maroto, M.C.; Schneider, R.; Baumes, R. Formation pathways of ethyl esters of branched short-chain fatty acids during wine aging. *Journal of Agricultural and Food Chemistry*, **2005**, 53, 3503-3509.

Drost, B.W.; Van Eerde, P.; Hoekstra, S.F.; Strating, J. Fatty acids and staling of beer. *Proceedings of the European Brewery Convention Congress*, **1971**, 451–458.

Duarte, I.; Barros, A.; Belton, P.S.; Righelato, R.; Spraul, M.; Humpfer, E.; Gil, A.M. High-resolution Nuclear Magnetic Resonance spectroscopy and multivariate analysis for the characterization of beer. *Journal of Agricultural and Food Chemistry*, **2002**, 50, 2475-2481.

Duarte, I.F.; Godejohann, M.; Braumann, U.; Spraul, M.; Gil, A.M. Application of NMR spectroscopy and LC-NMR/MS to the identification of carbohydrates in beer. *Journal of Agricultural and Food Chemistry*, **2003**, 51, 4847-4852.

Duarte, I.F.; Barros, A.; Almeida, C.; Spraul, M.; Gil, A.M. Multivariate analysis of NMR and FTIR data as a potential tool for the quality of beer. *Journal of Agricultural and Food Chemistry*, **2004**, 52, 1031-1038.

Duarte, I.F.; Goodfellow, B.J.; Gil, A.M. Characterization of mango juice by high-resolution NMR, hyphenated NMR, and diffusion-ordered spectroscopy. *Spectroscopy Letters*, **2005**, 38, 319-342.

Elipe, M.V.S. Advantages and disadvantages of nuclear magnetic resonance spectroscopy as a hyphenated technique. *Analytica Chimica Acta*, **2003**, 497, 1-25.

Erbe, T.;Bruckner, H. Chromatographic determination of amino acid enantiomers in beers and raw materials used for their manufacture, *Journal of Chromatography A*, **2000**, 881, 81-91.

Ernandes, J.R.; Williams, J.W.; Russell, I.; Stewart, G.G. Respiratory deficiency in brewing yeast: effects on fermentation, flocculation, and beer flavor components. *Journal of the American Society of Brewing Chemists*, **1993**, 51, 16-20, 1993.

Erny, G.L.; Rodrigues, J.E.A.; Gil, A.M.; Barros, A.S.; Esteves, V.I. Analysis of non-aromatic organic acids in beer by CE and direct detection mode with diode array detection. *Chromatographia*, **2009**, 70, 1737-1742.

Esteves, V.I.; Lima, S.S.F.; Lima, D.L.D.; Duarte, A.C. Using capillary electrophoresis for the determination of organic acids in Port wine. *Analytica Chimica Acta*, **2004**, 513, 163-167.

Evilia, R.F. Quantitative NMR spectroscopy. *Analytical Letters*, **2001**, 34, 2227-2236.

Exarchou, V.; Krucker, M.; van Beek, T.A.; Vervoot, J.; Gerothanassis, P.; Albert, K. LC-NMR coupling technology: recent advancements and applications in natural products analysis. *Magnetic Resonance in Chemistry*, **2005**, 43, 681-687.

Fan, T.W.M. Metabolite profiling by one- and two-dimensional NMR analysis of complex mixtures. *Progress in Nuclear Magnetic Resonance Spectroscopy*, **1996**, 28, 161-219.

Fan, T.W.M.; Lane, A.N. Structure-based profiling of metabolites and isotopomers by NMR. *Progress in Nuclear Magnetic Resonance Spectroscopy*, **2008**, 52, 69-117.

Faber, N.M.; Rajko, R. How to avoid over-fitting in multivariate calibration – the conventional validation app2oach and an alternative. *Analytica Chimica Acta*, **2007**, 595, 98-106.

Ferre, H.; Broberg, A.; Duus, J.O.; Thomsen, K.K. A novel type of arabinoxylan arabinofuranohydrolase isolated from germinated barley. *European Journal of Biochemistry*, **2000**, 267, 6633-6641.

Ferreira, V.; Jarauta, I.; Lopez, R.; Cacho, J. Quantitative determination of sotolon, maltol and free furaneol in wine by solid-phase extraction and gas-chromatography-ion-trap mass spectrometry. *Journal of Chromatography A*, **2003**, 1010, 95-103.

François, N.; Guyot-Declerck, C.; Hug, B.; Callemien, D.; Govaerts, B.; Collin, S. Beer astringency assessed by time–intensity and quantitative descriptive analysis: Influence of pH and accelerated aging. *Food Quality and Preference*, **2006**, 17, 445-452.

Franconi, J.M.; Naulet, N.; Martin, G.J. Natural factors of site-specific isotope fractionation of hydrogen in the fermentation products (ethanol and water) of barley. *Journal of Cereal Science*, **1989**, 10, 69-80.

Frazier, R.A; Ames, J.M.; Nursten, H.E. Background theory and principles of capillary electrophoresis. In *Capillary Electrophoresis for Food Analysis – Method Development*; RSC: Cambridge, 2000, 1-7.

Gardner, R. J.; McGuinness, J. D. Complex polyphenols in brewing: a critical survey. *Master Brewers Association of the Americas Technical Quarterly*, **1977**, *14*, 250-261.

Garrod, S.; Humpher, E; Connor, S.C.; Connelly, J.C.; Spraul, M.; Nicholson, J.K.; Holmes, E. High-resolution (1)H NMR and magic angle spinning NMR spectroscopic investigation of the biochemical effects of 2-bromoethanamine in intact renal and hepatic tissue. *Magnetic Resonance in Medicine*, **2001**, 45, 781–790.

Gijs, L.; Perpete, P.; Timmermans, A.; Collin, S. 3-Methylthiopropionaldehyde as precursor of dimethyltrisulfide in aged beers. *Journal of Agricultural and Food Chemistry*, **2000**, 48, 6196-6199.

Gijs, L.; Chevance, F.; Jerkovic, V.; Collin, S. How low pH can intensify α-damascenone and dimethyl trisulfide production through beer aging. *Journal of Agricultural and Food Chemistry*, **2002a**, 50, 5612-5616.

Gijs, L.; Collin, S. Effect of the reducing power of a beer on dimethyltrisulfide production during aging. *Journal of the American Society of Brewing Chemists*, **2002b**, 60, 68-70.

Gil, A.M.; Duarte, I.F.; Godejohann, M.; Braumann, U.; Maraschin, M.; Spraul, M. Characterization of the aromatic composition of some liquid foods by nuclear magnetic resonance spectrometry and liquid chromatography with nuclear magnetic resonance and mass spectrometric detection. *Analytica Chimica Acta*, **2003**, 488, 35-51.

Goldammer, T. *The Brewer's Handbook: the Complete Book to Brewing Beer*, 2nd Edition; Apex Publishers: Clifton, 2008.

Gomez-Carracedo, M. P.; Andrade, J. M.; Fernandez, E.; Prada, D.; Muniategui, S. Evaluation of the pure apple juice content in commercial apple beverages using FTMIR-ATR and potential curves. *Spectroscopy Letters*, **2004**, 37, 73–93.

Graça, G.; Duarte, I.F.; Barros, A.S.; Goodfellow, B.J.; Diaz, S.; Carreira, I.M.; Couceiro, A.B.; Galhano, E.; Gil, A.M. ^{1}H NMR based metabonomics of human amniotic fluid for the metabolic characterization of fetus malformations. *Journal of Proteome Research*, **2009**, 8, 4144-4150.

Grob, R.L. *Modern Practice of Gas Chromatography*, 3rd Edition; John Wiley & Sons: New York, 1995.

Guido, L.F.; Fortunato, N.A.; Rodrigues, J.A.; Barros, A.A. Voltammetric assay for the aging of beer. *Journal of Agricultural and Food Chemistry*, **2003**, 51, 3911-3915.

Guido, L.F.; Carneiro, J.R.; Santos, J.R.; Almeida, P.J.; Rodrigues, J.A.; Barros, A.A. Simultaneous determination of E-2-nonenal and β-damascenone in beer by reversed-phase liquid chromatography with UV detection. *Journal of Chromatography A*, **2004**, 1032, 17-22.

Guido, L.F.; Curto, A.; Boivin, P.; Benismall, N.; Gonçalves, C.; Barros, A.A. Predicting the organoleptic stability of beer from chemical data using multivariate analysis. *European Food Research and Technology*, **2007**, 226, 57-62.

Guyot-Declerck, C.; François, N.; Ritter, C.; Govaerts, B.; Collin, S. Influence of pH and ageing on beer organoleptic properties. A sensory analysis based on AEDA data. *Food Quality and Preference*, **2005**, 16, 157-162.

Hardwick, W.A. The properties of beer. In *Handbook of brewing*; Hardwick, W.A., Ed.; Marcel Dekker: New York, 1994, 551-586.

Hashimoto, N. Oxidation of higher alcohols by melanoidins in beer. *Journal of the Institute of Brewing*, **1972**, 78, 43–51.

Hashimoto, N.; Kuroiwa, Y. Proposed pathways for the formation of volatile aldehydes during storage of bottled beer. *Proceedings of the American Society of Brewing Chemists*, **1975**, 33, 104-111.

Hashimoto, N.; Eshima, T. Oxidative degradation of isohumulones in relation to flavour stability of beer. *Journal of the Institute of Brewing*, **1979**, 85, 136–140.

Holmes, E.; Nicholson, J. K.; Tranter, G. Metabonomic characterization of genetic variations in toxicological and metabolic responses using probabilistic neural networks. *Chemical Research Toxicology*, **2001**, 14, 182-191.

Horak, T.; Culik, J.; Cejka, P; Jurkova, M.; Kellner, V.; Dvorak, J.; Haskova, D. Analysis of free fatty acids in beer: comparison of solid-phase extraction, solid-phase microextraction, and stir bar sportive extraction. *Journal of Agricultural and Food Chemistry*, **2009**, 57, 11081-11085.

Hornsey, I.S. *A History of Beer and Brewing*; RSC: Cambridge, 2003.

Hughes, P.S.; Baxter, E.D. *Beer: Quality, Safety and Nutritional Aspects*; RSC: Cambridge, 2001.

http://www.chem.queensu.ca, consulted on 10th June, 2011.

http://www.perkinelmer.com, consulted on 15th May, 2011.

Intelmann, D.; Kummerlowe, G.; Haseleu, G.; Desmer, N.; Schulze, K.; Frohlich, R.; Franck, O.; Luy, B.; T. Hofmann, T. Structures of storage-induced transformation products of the beer's bitter principles, revealed by sophisticated NMR spectroscopic and LC-MS techniques. *Chemistry – a European Journal*, **2009**, 15, 13047-13058.

Intelmann, D.; Hoffmann, T. On the oxidation of bitter-tasting iso-α-acids in beer. *Journal of Agricultural and Food Chemistry*, **2010**, *58*, 5059-5067.

Irwin, A.J.; Barker, R.L.; Pipasts, P. The role of copper, oxygen, and polyphenols in beer flavor instability. *Journal of the American Society of Brewing Chemists*, **1991**, 49, 140–148.

Jacobsen, N.E. *NMR Spectroscopy Explained – simplified Theory, Applications and Examples for Organic Chemistry and Structural Biology*; John Wiley & Sons: New York, 2007.

Jaillais, B.; Pinto, R.; Barros, A.S.; Rutledge, D.N. Outer-product analysis (OPA) using PCA to study the influence of temperature on NIR spectra of water. *Vibrational Spectroscopy*, **2005**, 39, 50-58.

James, D.; Scott, S.M.; O'Hare, W.T.; Ali, Z.; Rowell, F.J. Classification of fresh edible oils using a coated piezoelectric sensor array-based electronic nose with soft computing approach for pattern recognition. *Transactions of the Institute of Measurement and Control*, **2004**, 26, 3-18.

Jennings, W.; Mittlefehldt, E.; Stermple, P. *Analytical Gas Chromatography*, 2nd Edition; Academic Press: San Diego, 1997.

Jiang, Z.; Liang, Q.; Wang, Y.; Zheng, X.; Pei, L.; Zhang, T.; Wang, Y.; Luo, G. Metabonomic study on women of reproductive age treated with nutritional intervention: screening potential biomarkers related to neural tube defects occurence. *Biomedical Chromatography*, **2010**, 25, 767-774.

Jodelet, A.; Rigby, N.M.; Colquhoun, I.J. Separation and NMR structural characterisation of singly branched α-dextrins which differ in the location of the branch point. *Carbohydrate Research*, **1998**, 312, 139-151.

Jollife, I.T. Principal Component Analysis; Springer: New York, 1986.

Jonkova, G.; Petkova, N. Effect of some technological factors on the content of acetaldehyde in beer. *Journal of the University of Chemical Technology and Mettalurgy*, **2011**, 46, 57-60.

Kadkhodaie, M.; Rivas, O.; Tan, M.; Mohebbi, A.; Shaka, A.J. Broad-band homonuclear cross polarization using flip-flop spectroscopy. *Journal of Magnetic Resonance*, **1991**, 91, 437-443.

Kaneda, H.; Kano, Y.; Koshino, S.; Ohyanishiguchi, H. Behavior and role of iron ions in beer deterioration. *Journal of Agricultural and Food Chemistry*, **1992**, 40, 2102–2107.
Kaneda, H.; Kobayashi, M.; Takashio, M.; Tamaki, T.; Shinotsuka, K. Beer staling mechanism. *Master Brewers Association of the Americas Technical Quarterly*, **1999**, 36, 41–47.

Karasek, F.W.; Clement, R.E. *Basic Gas Chromatography-Mass Spectrometry: principles and techniques*; Elsevier: Amsterdam, 1988.

Katajamaa, M.; Oresic, M. Data processing for mass spectrometry-based metabolomics. *Journal of Chromatography A*, **2007**, 1158, 318-328.

Keniry, M. A.; Sanctuary, B. C. Experimental verification of composite inversion pulses obtained by series expansion of the offset angle, *Journal of Magnetic Resonance*, **1992**, 97, 382–384.

Khatib, A.; Wilson, E.G.; Kim, H.K.; Lefeber, A.W.M.; Erkelens, C.; Choi, Y.H.; Verpoorte, R. Application of two-dimensional J-resolved nuclear magnetic resonance spectroscopy to differentiation of beer. *Analytica Chimica Acta*, **2006**, 559, 264-270.

Khatib, A.; Wilson, E.G.; Kim, H.K.; Supardi, M.; Choi, Y.H.; Verpoorte, R. NMR assignment of iso-alpha-acids from isomerised extracts of *Humulus lupulus* L. Cones. *Phytochemical Analysis*, **2007**, 18, 371-277.

Kim, M.O.; Baltes, W. On the role of 2,3-dihydro-3,5-dihydroxy-6-methyl-4(H)-pyran-4-one in the Maillard reaction. *Journal of Agricultural and Food Chemistry*, **1996**, 44, 282-289.

Kind, T.; Fiehn, O. Advances in structure elucidation of small molecules using mass spectrometry. *Bionalytical Reviews*, **2010**, 2, 23-60.

Kirwan, G.M.; Clark, S.; Barnett, N.W.; Niere, J.O.; Adams, M.J. Generalised 2D-correlation NMR analysis of a wine fermentation. *Analytica Chimica Acta*, **2008**, 629, 128-135.

Kirwan, G.M.; Coffey, V.G.; Niere, J.O.; Hawley, J.A.; Adams, M.J. Spectroscopic correlation analysis of NMR-based metabonomics in exercise science. *Analytica Chimica Acta*, **2009**, 652, 173-179.

Klampfl, C.W. Analysis of organic acids and inorganic anions in different types of beer using capillary zone electrophoresis. *Journal of Agricultural and Food Chemistry*, **1999**, 47, 987-990.

Klein, E. *A comprehensive etymological dictionary of the English language*; Elsevier Scientific Publishing: Amsterdam, 1971.

Kocherginsky, N.M.; Kostetski, Y.Y.; Smirnov, A.I. Antioxidant pool in beer and kinetics of EPR spin-trapping. *Journal of Agricultural and Food Chemistry*, **2005**, 53, 6870-6876.

Kuchel, L.; Brody, A.L.; Wicker, L. Oxygen and its reactions in beer; *Packaging Technology and Science*, **2006**, 19, 25-32.

Lachenmeier, D.W.; Franck, W.; Humpfer, E.; Schafer, H.; Keller, S.; Mortter, M.; Spraul, M. Quality control of beer using high-resolution nuclear magnetic resonance spectroscopy and multivariate analysis. *European Food Research and Technology*, **2005**, 220, 215-221.

Lambert, J.B.; Mazzola, E.P. *Nuclear Magnetic Resonance Spectroscopy: An introduction to principles, applications, and experimental methods*; Prentice Hill; New Jersey, 2003.

Larsen, F.H.; van den Berg, F.; Engelsen, S.B. An exploratory chemometric study of ^{1}H NMR spectra of table wines. *Journal of Chemometrics*, **2006**, 20, 198-208.

Le Gall, G.; Puaud, M.; Colquhoun, I.J. Discrimination between orange juice and pulp wash by ^{1}H nuclear magnetic resonance spectroscopy: identification of marker compounds. *Journal of Agricultural and Food Chemistry*, **2001**, 49, 580-588.

Lee, W.J.; Prentice, N. Utilization of nucleosides and nucleobases by the lager yeast Saccharomyces carlsbergenis. *Journal of the American Society of Brewing Chemists*, **1987**, 45, 128-131.

Leick, G. *The A to Z of Mesopotamia*; Scarecrow Press: Plymouth, 2010.

Lemon, J. Plotrix: a package in the red light district of R. *R-news*, **2010**, 6, 8-12.

Lermusieau, G.; Noel, S.; Liegeois, C.; Collin, S. Nonoxidative mechanism for development of trans-2-nonenal in beer. *Journal of the American Society of Brewing Chemists*, **1999**, 57, 29-33.

Lermusieau, G.; Liegeois, C.; Collin, S. Reducing power of hop cultivars and beer ageing. *Food Chemistry*, **2001**, 72, 413-418.

Liegeois, C.; Meurens, N.; Badot, C.; Collin, S. Release of deuterated (E)-2-nonenal during beer aging from labeled precursors synthesized before boiling. *Journal of Agricultural and Food Chemistry*, **2002**, 50, 7634–7638.

Lindon, J.C.; Nicholson, J.K.; Holmes, E. *Handbook of Metabonomics and Metabolomics*; Elsevier: Amsterdam, 2007.

Lindon, J.C.; Nicholson, J.K. Spectroscopic and statistical techniques for information recovery in metabonomics and metabolomics. *Annual Review of Analytical Chemistry*, **2008**, 1, 45-69.

Liu, M.; Nicholson, J.K.; Lindon, J.C. High resolution diffusion and relaxation edited one and two-dimensional 1H NMR spectroscopy of biological fluids. *Analytical Chemistry*, **1996**, 68, 3370–3376.

Liu, J.; Dong, J.; Chen, J.; Gu, G. Multivariate modeling of aging in bottled lager beer by principal component analysis and multiple regression methods. *Journal of Agricultural and Food Chemistry*, **2008**, 56, 7106-7112.

Lopez-Rituerto, E.; Cabredo, S.; Lopez, M.; Avenoza, A.; Busto, J.H.; Peregrina, J.M. A thorough study on the use of quantitative ^{1}H NMR in rioja red wine fermentation processes. *Journal of Agricultural and Food Chemistry*, **2009**, 57, 2112-2118.

Madigan, D.; Perez, A.; Clements, M. Furanic aldehyde analysis by HPLC as a method to determine heat-induced flavor damage to beer. *Journal of the American Society of Brewing Chemists*, **1998**, 56, 146–151.

Maher, A.D.; Cysique, L.A.; Brew, B.J.; Rae, C.D. Statistical integration of ^{1}H NMR and MRS data from different biofluids and tissues enhances recovery of biological information from individuals with HIV-1 infection. *Journal of Proteome Research*, **2011**, 10, 1737-1745.

Malfliet, S.; Goiris, K.; Aerts, G.; De Cooman, L. Analytical-sensory determination of potential flavor deficiencies of light beers. *Journal of the Institute of Brewing*, **2009**, 115, 49-63.

Mannina, L.; Patumi, M.; Proietti, N.; Bassi, D.; Segre, A.L. Geographical characterization of Italian extra virgin olive oils using high-field ^{1}H NMR spectroscopy. *Journal of Agricultural and Food Chemistry*, **2001**, 49, 2687-2696.

Mannina, L.; Dugo, G.; Salvo, F.; Cicero, L.; Ansanelli, G.; Calgagni, C.; Segre, A. Study of the cultivar-composition relationship in Sicilian olive oils by GC, NMR, and statistical methods. *Journal of Agricultural and Food Chemistry*, **2003**, 51, 120-127.

Martin, G.J.; Benbernou, M.; Lantier, F. Application of site-specific natural isotope fractionation (SNIF-NMR) of hydrogen to the characterization of europeans beers. *Journal of the Institute of Brewing*, **1985**, 91, 242-249.

Matisova, E.; Domotorova, M. Fast gas chromatography and its use in trace analysis – review. *Journal of Chromatography A*, **2003**, 1000, 199-221.

Mato, I.; Suarez-Luque, S.; Huidobro, J.F. Simple determination of main organic acids in grape juice and wine by using capillary zone electrophoresis with direct UV detection. *Food Chemistry*, **2007**, 102, 104-112.

McMaster, M.C. *GC/MS – A Practical User's Guide*; John Wiley & Sons: New York, 2008.

Meany, J.E. Lactate dehydrogenase catalysis: roles of keto, hydrated, and enol pyruvate. *Journal of Chemical Education*, 2**007**, 84, 1520-1523.

Meilgaard, M. Stale flavor carbonyls in brewing. *Brewers Digest*, **1972**, 47, 48-57.

Moco, S.; Forshed, J.; De Vos, R.C.H:; Bino, R.J.; Vervoot, J. Intra- and inter-metabolite correlation spectroscopy of tomato metabolomics data obtained by liquid chromatography-mass spectrometry and nuclear magnetic resonance. *Metabolomics*, **2008**, 4, 202-215.

Montanari, L.; Perretti, G.; Natella, F.; Guidi, A.; Fantozzi, P. Organic and phenolic acids in beer; *Lebensmittel-Wissenschaft und-Tecnologie*, **1999**, 32, 535-539.
Narziss, L. *Abriss der Bierbrauerei*; Wiley-VCH: Weinheim, 2005.

Nicholson, J.K.; Foxall, P.J.D.; Spraul, M.; Farrant, R.D.; Lindon, J.C. 750 MHz 1H and 1H-13C NMR spectroscopy of human blood plasma. *Analytical Chemistry*, **1995**, 67, 793-811.

Nicholson, J.K.; Lindon, J.C.; Holmes. E. 'Metabonomics': Understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica*, **1999**, 29, 1181-1189.

Niessen, W.M.A. *Current Practice of Gas Chromatography-Mass Spectrometry*; Marcel Dekker: New York, 2001.

Ninonuevo, M.R.; Park, Y.; Yin, H.; Zhang, J.; Ward, R.E.; Clowers, B H.; German, J.B.; Freeman, S.L.; Killeen, K.; Grimm, R.; Lebrilla, C.B. A strategy for annotating the human milk glycome. *Journal of Agriculture and Food Chemistry*, **2006**, 54, 7471-7480.

Noda, I. Generalized two-dimensional correlation method applicable to infrared, raman, and other types of spectroscopy. *Applied Spectroscopy*, **1993**, 47, 1329-1336.

Noda, I.; Ozaki, Y. *Two-dimensional Correlation Spectroscopy – Applications in Vibrational and Optical Spectroscopy*, John Wiley and Sons, Ltd, Chichester, 2004.

Noel, S.; Metais, N.; Bonte, S.; Bodart, E.; Peladan, F.; Dupire, S.; Collin, S. The use of oxygen 18 in appraising the impact of oxidation process during beer storage. *Journal of the Institute of Brewing*, **1999**, 105, 269-274.

Nord, L.I.; Sorensen, S.B.; Duus, J.O. Characterization of reduced iso-α-acids derived from hops (*Humulus lupulus*) by NMR. *Magnetic Resonance in Chemistry*, **2003**, 41, 660-670.

Nord. L.I.; Vaag, P.; Duus, J.O. Quantification of organic acids and amino acids in beer by ^{1}H NMR spectroscopy. *Analytical Chemistry*, **2004**, 76, 4790-4798.

Ogrinc, N.; Kosir, I.J.; Kocjancic, M.; Kidric, J. Determination of authenticy, regional origin, and vintage of Slovenian wines using a combination of IRMS and SNIF-NMR analyses. *Journal of Agricultural and Food Chemistry*, **2001**, 49, 1432-1440.

Palmer, G.H. Barley and malt. In *Handbook of Brewing*, 2nd Edition; Priest, F.G.; Stewart, G.G. Eds.; CRC Press: Boca Raton, 2006.

Pavia, D.L.; Lampman, G.M.; Kriz, G.S.; Vyvyan, J.R. *Introduction to Spectroscopy*, 4th Edition; Brooks/Cole: Belmont, 2009.

Peppard, T.L.; Halsey, S.A. The occurrence of 2 geometrical-isomers of 2,4,5-trimethyl-1,3-dioxolane in beer. *Journal of the Institute of Brewing*, **1982**, 88, 309–312.

Pereira, G.E.; Gaudillere, J.P.; Van Leeuwen, C.; Hilbert, G.; Maucourt, M; Deborde, C.; Moing, A.; Rolin, D. 1H NMR metabolic profiling of wines from three cultivars, three soil types and two contrasting vintages. *Journal International des Sciences de la Vigne et du Vin*, **2007**, 41, 103-109.

Pinho, O.; Ferreira, I.M.L.V.O.; Santos, L.H.M.L.M. Method optimization by solid-phase microextraction in combination with gas chromatography with mass spectrometry for analysis of beer volatile fraction. *Journal of Chromatography A*, **2009**, 1121, 145-153.

Powell, C.D.; Quain, D.E.; Smart, K.A. The impact of brewing yeast cell age on fermentation performance, attenuation and flocculation. *FEMS Yeast Research*, **2003**, 3, 149-157.

Priest, F.G.; Stewart, G.G. *Handbook of Brewing*, 2nd Edition; CRC Press: Boca Raton, 2006.

Pusecker, K.; Albert, K. Bayer, E. Investigation of hop and beer bitter acids by coupling of high-performance liquid chromatography to nuclear magnetic resonance spectroscopy. *Journal of Chromatography A*, **1999**, 836, 245-252.

Reid, L.M.; O'Donnell, C.P.; Downey, G. Recent technological advances for the determination of food authenticity. *Trends in Food Science and Technology*, **2006**, 17, 344-353.

Rezzi, S.; Ramadan, Z.; Martin, F.P.J.; Fay, L.B.; Van Bladeren, P.; Lindon, J.C.; Nicholson, J.K.; Kochhar, S. Human metabolic phenotypes link directly to specific dietary preferences in healthy individuals. *Journal of Proteome Research*, **2007**, 6, 4469-4477.

Rizzo, V.; Pinciroli, V. Quantitative NMR in synthetic and combinatorial chemistry. *Journal of Pharmaceutical and Biomedical Analysis*, **2005**, 38, 851-857

Rocha, S.M.; Coutinho, P.; Barros, A.; Delgadillo, I.; Coimbra, M.A. Rapid tool for distinction of wines based on the global volatile signature. *Journal of Chromatography A*, **2006**, 1114, 188-197.

Rood, D. *The Troubleshooting and Maintance Guide for Gas Chromatographers*. Wiley-VCH: Weinheim, 2007.

Ross, K.L.; Tu, T.T.; Smith, S.; Dalluge, J.J. Profiling of organic acids during fermentatation by ultraperformance liquid chromatography-tandem mass spectrometry. *Analytical Chemistry*, **2007**, 79, 4840-4844.

Saison, D.; De Schutter, D.P.; Delvaux, F.; Delvaux, F.R. Optimisation of a complete method for the analysis of volatiles involved in the flavor stability of beer by solid-phase microextraction in combination with gas chromatography and mass spectrometry. *Journal of Chromatography A*, **2008**, 1190, 342-349.

Rutledge, D.N.; Bouveresse, D.J.R. Multi-way analysis of outer product arrays using PARAFAC. *Chemometrics and Intelligent Laboratory Systems*, **2007**, 85, 170-178.
Saison, D.; De Schutter, D.P.; Uyttenhove, B.; Delvaux, F.; Delvaux, F.R. Contribution of staling compounds to the aged flavour of lager beer by studying their flavour thresholds. *Food Chemistry*, **2009**, 114, 1206-1215.

Saison, D.; De Schutter, D.P.; Vanbeneden, N.; Daenen, L.; Delvaux, F.; Delvaux, F.R. Decrease of aged beer aroma by the reducing activity of brewing yeast. *Journal of Agricultural and Food Chemistry*, **2010**, 58, 3107-3115.

Savel, J.; Kosin, P.; Broz, A. Anaerobic and aerobic beer aging. *Czech Journal of Food Science*, **2010**, 28, 18-26.

Savitzky, A.; Golay, M.J.E. Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, **1964**, 36, 1627-1639.

Shaykhutdinov, R.A.; MacInnis, G.D.; Dowlatabadi, R.; Weljie, A.M.; Vogel, H.J. Quantitative analysis metabolite concentrations in human urine samples using 13C{1H} NMR spectroscopy. *Metabolomics*, **2009**, 5, 307-317.

Shimizu, A.; Ikeguchi, M.; Sugai, S. Appropriateness of DSS and TSP as internal references for 1H NMR studies of molten globule proteins in aqueous media. *Journal of Biomolecular NMR*, **1994**, 4, 859-862.

Silva, G.A.; Augusto, F.; Poppi, R.J. Exploratory analysis of the volatile profile of beers by HS-SPME-GC. *Food Chemistry*, **2008**, 111, 1057-1063.

Silva Ferreira, A. C.; Guedes de Pinho, P.; Rodrigues, P.; Hogg, T. Kinetics of oxidative degradation of white wines and how they are affected by selected technological parameters. *Journal of Agricultural and Food Chemistry*, **2002**, *50*, 5919-5924.

Soares da Costa, M.; Alves, C.G.; Ferreira, A.; Ibsen, C.; Guedes de Pinho, P.; Silva Ferreira, A.C. Further insights into the role of methional and phenylacetaldehyde in lager beer flavor stability. *Journal of Agricultural and Food Chemistry*, **2004**, 52, 7911-7917.

Solanky, K.S.; Bailey, N.J.; Beckwith-Hall, B.M.; Bingham, S.; Davis, A.; Holmes, E.; Nicholson, J.K.; Cassidy, A. Biofluid ^{1}H NMR-based metabonomic techniques in nutrition research metabolic effects of dietary isoflavones in humans. *Journal of Nutritional Biochemistry*, **2005**, 16, 236-244.

Son, H.S.; Hwang, G.S.; Kim, K.M.; Ahn, H.J.; Park, W.M.; Van Den Berg, F.; Hong, Y.S.; Lee, C.H. Metabolomic studies on geopraphical grapes and their wines using 1H NMR analysis coupled with multivariate statistics. *Journal of Agricultural and Food Chemistry*, **2009**, 57, 1481-1490.

Spraul, M.; Schutz, B.; Humpfer, E.; Mortter, M.; Schafer, H.; Koswig, S.; Rinke, P. Mixture analysis by NMR as applied to fruit juice quality control. *Magnetic Resonance Chemistry*, **2009**, 47, 130-137.

Stafford, D. Chromatography. In *Principles of Forensic Toxicology*, 2nd Edition, Levine, B. Ed.; AACC Press: Washington, 2003, 89-116.
Stuart, B. *Infrared Spectroscopy: Fundamentals and Applications*; John Wiley & Sons: Chichester, 2004.

Thum, B.; Miedaner, H.; Narziss, L.; Back, W. Bildung von "*Alterungscarbonylen*" - mögliche Mechanismen un Bedeutung bei der Bierlagerung. *Proceedings of the European Brewery Convention Congress*, **1995**, 491-498.

Tiziani, S.; Schwartz, S.J.; Vodovotz, Y. Profiling of carotenoids in tomato juice by one- and two-dimensional NMR. *Journal of Agricultural and Food Chemistry*, **2006**, 54, 6094-6100.

Tressl, R.; Helak, B.; Rewicki, D. Maltoxazin, eine tricyclishe verbindung aus gerstenmalz. *Helvetica Chimica Acta*, **1982**, 65, 483-489.

Uchida, M.; Ono, M.; Nagami, K. Carbohydrates in brewing. I. Determination of fermentable sugars and oligosaccharides in wort and beer by partition high-performance liquid chromatography. *Journal of the American Society of Brewing Chemists*, **1991**, 49, 65-73.

Vanderhaegen, B.; Neven, H.; Coghe, S.; Verstrepen, K.J.; Verachtert, H.; Derdelinckx, G. Evolution of chemical and sensory properties during aging of top-fermented beer. *Journal of Agricultural and Food Chemistry*, **2003**, 51, 6782-6790.

Vanderhaegen, B.; Neven, H.; Daenen, L.; Verstrepen, K.J.; Verachtert, H.; Derdelinckx, G. Furfuryl ethyl ether: Important aging flavor and a new marker for the storage conditions of beer. *Journal of Agricultural and Food Chemistry*, **2004a**, 52, 1661-1668.

Vanderhaegen, B.; Neven, H.; Verstrepen, K.J.; Delvaux, F.R.; Verachtert, H.; Derdelinckx, G. Influence of the brewing process on furfuryl ethyl ether formation during beer aging. *Journal of Agricultural and Food Chemistry*, **2004b**, 52, 6755-6764.

Vanderhaegen, B.; Neven, H.; Verachtert, H.; Derdelinckx, G. The chemistry of beer aging – a critical review. *Food Chemistry*, **2006**, 95, 357-381.

Vanderhaegen, B.; Delvaux, F.; Daenen, L.; Verachtert, H.; Delvaux, F.R. Aging characteristics of different beer types. *Food Chemistry*, **2007**, 103, 404- 412.

Varmuza, K.; Steiner, I.; Glinsner, T.; Klein, H. Chemometric evaluation of concentration profiles from compounds relevant in beer ageing. *European Food Research and Technology*, **2002**, 215, 235-239.

Verzele, M.; De Keuleire, D. *Chemistry and Analysis of Hop and Beer Bitter Acids*; Elsevier: Amesterdam, 1991.

Viereck, N.; Norgaard, L.; Bro, R.; Engelsen, S.B. Chemometric Analysis of NMR Data. In *Modern Magnetic Resonance*, Volumes 1-3; Webb, G.W. Ed.; Springer, 2008, 1833-1843.

Viggiani, L.; Morelli, M.A.C. Characterization of wines by nuclear magnetic resonance: a work study on wines from the Basilicata region in Italy. *Journal of Agricultural and Food Chemistry*, **2008**, 56, 8273-8279.

Vigli, G.; Philippidis, A.; Spyros, A.; Dais, P. Classification of edible oils by employing P-31 and H-1 NMR spectroscopy in combination with multivariate statistical analysis. A proposal for the detection of seed oil adulteration in virgin olive oils. *Journal of Agricultural and Food Chemistry*, **2003**, 51, 5715-5722.

Vinogradov, E.; Bock, K.; Structural determination of some new oligosaccharides and analysis of the branching pattern of isomaltooligosaccharides from beer. *Carbohydrate Research*, **1998**, 309, 57-64.

Weljie, A.M.; Newton, J.; Mercier, P.; Carlson, E.; Slupsky, C.M. Targeted Profiling: quantitative analysis of 1H NMR metabolomics data. *Analytical Chemistry*, **2006**, 78, 4430-4442.

Westerhuis, J.A.; Hoefsloot, H.C.J.; Smit, S.; Vis, D.J.; Smilde, A.K.; van Velzen, J.J.; van Duijnhoven, J.P.M; van Dorsten, F.A. Assessment of PLSDA cross validation. *Metabolomics*, **2008**, 4, 81-89.

Willaert, R.; Nedovic, V.A. Primary beer fermentation by immobilised yeast – a review on flavor formation and control strategies. *Journal of Chemical Technology and Biotechnology*, **2006**, 81, 1353-1367.

Williams, R.S.; Wagner, H.P. The isolation and identification of new staling related compounds form beer. *Journal of the American Society of Brewing Chemists*, **1978**, 36, 27-31.

Wold, S.; Sjostrom, M.; Eriksson, L. PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, **2001**, 58, 109-130.

Woonton, B.W.; Sherkat, F.; Maharjan, P. The influence of barley storage on respiration and glucose-6-phophate dehydrogenase during malt. *Journal of the Institute of Brewing* **2005**, 111, 388-395.

Yan, Z.; Xingde, Z.; Weijun, N. Simultaneous determination of carbohydrates and organic acids in beer and wine by ion chromatography. *Mikrochimica Acta*, **1997**, 127, 189-194.

Yaylayan, V.A; Mandeville, S. Mechanistic pathway for the formation of maltoxazine from intact 1-[(‘2-carboxyl)pyrrolidinyl]-1-deoxy-$_D$-fructose (Amadori-proline). *Journal of Agricultural and Food Chemistry*, **1994**, 42, 1841-1844.

Zhao, H.; Chen, W.; Lu, J.; Zhao, M. Phenolic profiles and antioxidant activities of commercial beers. *Food Chemistry*, **2010**, 119, 1150-1158.

Zufall, C.; Tyrell, T. The influence of heavy metal ions on beer flavour stability. *Journal of the Institute of Brewing*, **2008**, 114, 134-142.